HONI SOIT QUI MAL Y PENSE
DIEU ET MON DROIT

AGRICULTURAL RESEARCH INSTITUTE
PUSA

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A

CONTAINING PAPERS OF A MATHEMATICAL AND
PHYSICAL CHARACTER.

VOL. XCIX.

LONDON:
PRINTED FOR THE ROYAL SOCIETY AND SOLD BY
HARRISON AND SONS, LTD., ST. MARTIN'S LANE,
PRINTERS IN ORDINARY TO HIS MAJESTY.

SEPTEMBER, 1921

LONDON:
HARRISON AND SONS, LTD., PRINTERS IN ORDINARY TO HIS MAJESTY,
ST. MARTIN'S LANE.

CONTENTS.

SERIES A. VOL. XCIX.

PAGE

Minutes of Meetings, January 20 to June 30, 1921* .... i–xviii

No. A 696.—April 6, 1921.

On Electro-Crystalline Properties as Conditioned by Atomic Lattices. By Sir Joseph Larmor, F.R.S., Cambridge 1

The Colour of the Light from the Night Sky. By Lord Rayleigh, F.R.S. 10

Linear Transformations and Functions of Positive Type. By J. Mercer, M.A., D.Sc., Fellow and Lecturer of Christ's College, Cambridge. Communicated by Prof. E. W. Hobson, F.R.S. 19

The Dissipation of Energy in Permanent Ocean Currents, with some Relations between Salinities, Temperatures and Currents. By R. O. Street, M.A., M.Sc., Fellow of St. John's College, Cambridge, Lecturer in Mathematics and Mathematical Hydrography in the University of Liverpool. Communicated by Sir Joseph Larmor, F.R.S. 39

The Constitution of the Alloys of Aluminium, Copper, and Zinc containing High Percentages of Zinc. By John L. Haughton, D.Sc., F.Inst.P., and Kathleen E. Bingham, M.Sc. (both of the National Physical Laboratory). Communicated by Dr. W. Rosenhain, F.R.S. (Plates 1 and 2) 47

The Vacuum Arc Spectra of Sodium and Potassium. By Snehamoy Datta, M.Sc. (Calcutta), D.I.C., Research Student, Imperial College of Science. Communicated by Prof. A. Fowler, F.R.S. 69

On Spectrophotometry in the Visible and Ultra-violet Spectrum. By T. R. Merton, M.A., D.Sc., F.R.S., Professor of Spectroscopy in the University of Oxford 78

A Superior Limit to the Age of the Earth's Crust. By Henry Norris Russell, Professor of Astronomy in Princeton University. Communicated by Mr. J. H. Jeans, Sec.R.S. 84

No. A 697.—May 2, 1921.

Discussion on Isotopes. Opened by Sir J. J. Thomson, O.M., F.R.S. 87

A Generalisation of Weyl's Theory of the Electromagnetic and Gravitational Fields. By Prof. A. S. Eddington, F.R.S. 104

The Field of an Electron on Einstein's Theory of Gravitation. By G. B. Jeffery, M.A., D.Sc., Fellow of University College, London. Communicated by Prof. L. N. G. Filon, F.R.S. 123

---

* The Minutes of the January Meetings were issued with No. A 694, which No. formed part of the preceding volume.

PAGE

On a Physical Theory of Stellar Spectra. By M. N. Saha, D.Sc., Lecturer in Physics and Applied Mathematics, Calcutta University. Communicated by Prof. A. Fowler, F.R.S. 135

On the Catalytic Dehydrogenation of Alcohols. By Eric Keightley Rideal. Communicated by Sir Wm. Pope, F.R.S. 153

On the Design of Diaphragms Capable of Continuous Tuning. By Louis Vessot King, M.A. (Cantab.), D.Sc., Macdonald Professor of Physics, McGill University, Montreal. Communicated by Prof. A. S. Eve, F.R.S. 163

A Method of Analysing Galvanometer Records. By W. Hartree, M.A., and A. V. Hill, Sc.D., F.R.S. 172

No. A 698.—June 1, 1921.

On the Influence of Low Temperatures on the Magnetic Properties of Alloys of Iron with Nickel and Manganese. By Prof. Kamerlingh Onnes, For. Mem. R.S., Sir Robert A. Hadfield, Bart., F.R.S., and Dr. H. R. Woltjer 174

The Hardness of Solid Solutions. By Walter Rosenhain, B.A., D.Sc., F.R.S. (of the National Physical Laboratory) 196

The Influence of Physical Conditions on the Velocity of Decomposition of certain Crystalline Solids. By C. N. Hinshelwood, B.A., Fellow of Balliol College, and E. J. Bowen, B.A., Balliol College. Communicated by Prof. F. A. Lindemann, F.R.S. 203

Heats of Combustion and Formation of Nitro-Compounds. Part I.—Benzene, Toluene, Phenol and Methylaniline Series. By W. E. Garner and C. L. Abernethy. Communicated by Sir Robert Robertson, F.R.S. 213

Researches upon Brown Coals and Lignites. Part I.—Heat Treatment at Temperatures below 400° C. as a Possible Method for Enhancing their Fuel Values. By William A. Bone, D.Sc., F.R.S. 236

On the Transformation of Integrals. By W. H. Young, Sc.D., F.R.S., Professor of Mathematics in the University of Wales 252

On the Moving Striations in a Neon Tube. By T. Kikuchi, Cavendish Laboratory. Communicated by Sir E. Rutherford, F.R.S. 257

The Magnetic Spectrum of the β-Rays Excited by γ-Rays. By C. D. Ellis, B.A., Coutts-Trotter Student of Trinity College, Cambridge. Communicated by Sir Ernest Rutherford, F.R.S. 261

No. A 699.—July 1, 1921.

The Vibrations of a Spinning Disk. By H. Lamb, F.R.S., and R. V. Southwell ... 272

A Liquid Oxygen Vaporiser. By Edgar A. Griffiths. Communicated by the Tyndall Mining Bequest Committee 281

A Simple Apparatus for Approximate Harmonic Analysis and for Periodicity Measurements. By W. Lawrence Balls, M.A., Sc.D. (Cantab.), Late Fellow of St. John's College, Cambridge. Communicated by Dr. F. F. Blackman, F.R.S. 283

PAGE

Multenions and Differential Invariants. By Alex. McAulay, M.A., Professor of Mathematics, University of Tasmania. Communicated by Mr. W. B. Hardy, Sec.R.S. 292

A New Form of Wehnelt Interrupter. By F. H. Newman, M.Sc., A.R.C.Sc. Communicated by Lord Rayleigh, F.R.S. (Plates 3 and 4) 324

On the Viscosity and Molecular Dimensions of Gaseous Cyanogen. By A. O. Rankine, D.Sc., Professor of Physics in the Imperial College of Science and Technology. Communicated by Prof. H. L. Callendar, F.R.S. 331

The Properties and Molecular Structure of Thin Films of Palmitic Acid on Water. Part I. By N. K. Adam, M.A., Fellow of Trinity College, Cambridge. Communicated by W. B. Hardy, Sec. R.S. 336

The Total Heat of Liquid Carbonic Acid. By C. Frewen Jenkin, C.B.E., M.Inst.C.E., and D. N. Shorthose, M.A. Communicated by Sir Alfred Ewing, K.C.B., F.R.S. 352

No. A 700.—August 2, 1921.

The Time Interval between Absorption and Emission of Light in Fluorescence. By R. W. Wood, For.Mem.R.S., Johns Hopkins University, Baltimore. (Plate 5) 362

A Study of the Glow of Phosphorus.—Periodic Luminosity and Action of Inhibiting Substances. By Lord Rayleigh, F.R.S. 372

Some Experiments on Thermal Diffusion. By T. L. Ibbs, M.C., B.Sc., Assistant Lecturer in Physics, University of Birmingham. With a Note on the Experiments by Prof. S. Chapman, F.R.S. Communicated by Prof. S. W. J. Smith, F.R.S. 385

The Dynamics of Revolving Fluid on a Rotating Earth. By Capt. D. Brunt, M.A., B.Sc. Communicated by Sir Napier Shaw, F.R.S. 397

Some Experiments on the Catalytic Reduction of Ethylene to Ethane. By Dorothy Muriel Palmer, Girton College, Cambridge, and William George Palmer, St. John's College, Cambridge. Communicated by Sir Wm. Pope, F.R.S. 402

The Catalytic Activity of Copper.—Part II. By William George Palmer, St. John's College, Cambridge. Communicated by Sir Wm. Pope, F.R.S. 412

A Reciprocating Expansion Apparatus for Detecting Ionising Rays. By Takeo Shimizu. Communicated by C. T. R. Wilson, F.R.S. 425

A Preliminary Note on Branched α-Ray Tracks. By Takeo Shimizu. Communicated by Sir E. Rutherford, F.R.S. 432

On the Spectra of the Alkaline Earth Fluorides and their Relation to each other. By Snehamoy Datta, M.Sc. (Calcutta), Research Student, Imperial College. Communicated by Prof. A. Fowler, F.R.S. (Plate 6) 436

PAGE

No. A 701.—September 1, 1921.

The Absorption of Light by Elements in a State of Vapour: The Halogens. By Sir J. J. Dobbie, F.R.S., and J. J. Fox, O.B.E., D.Sc. (Plate 7) 456

A Simple Extension of Fourier's Integral Theorem and some Physical Applications, in particular to the Theory of Quanta. By R. H. Fowler. Communicated by Prof. Sir E. Rutherford, F.R.S. 462

Some Problems connected with Evaporation from Large Expanses of Water. By M. A. Giblett, M.Sc. Communicated by Dr. G. C. Simpson, F.R.S. 472

A Study of Catalytic Actions at Solid Surfaces. VI.—Surface Area and Specific Nature of a Catalyst: two Independent Factors controlling the Resultant Activity. By E. F. Armstrong, D.Sc., F.R.S., and T. P. Hilditch, D.Sc. 490

Anomalous Liesegang Stratifications produced by the Action of Light. By Emil Hatschek. Communicated by Prof. A. W. Porter, F.R.S. (Plate 8) 496

On the Diffraction of Light incident at nearly the Critical Angle on the Boundary between two Media. By Brojendra Nath Chakravarty, M.Sc., Assistant to the Palit Professor of Physics, University of Calcutta. Communicated by Prof. J. W. Nicholson, F.R.S. (Plate 9) 503

Obituary Notices of Fellows Deceased:—

Sir W. de W. Abney (with portrait) i
Robert Bellamy Clifton vi
Sir Lazarus Fletcher (with portrait) ix
Srinivasa Ramanujan xiii
Woldemar Voigt xxix

Index xxxi

PROCEEDINGS OF

THE ROYAL SOCIETY.

Section A.—Mathematical and Physical Sciences.

On Electro-Crystalline Properties as Conditioned by Atomic Lattices.

By Sir Joseph Larmor, F.R.S., Cambridge.

(Received January 6, 1921.)

The method of analysis of crystal lattices by ultra-optical spectra, introduced and applied by W. L. and W. H. Bragg, has led to the view that the ultimate crystalline element is (or may be) the atom[*] and not the chemical molecule: so that each atom, presumably with its ionic electric charge, often stands in identically the same relation to a number of surrounding atoms of the other kind with opposite charge, and therefore cannot be associated physically with any one of them. The alternative view, previously perhaps the more natural on the physical as distinct from the tactical side, was to regard the chemical molecule as the crystalline unit: its associated positive and negative atomic ions might then form an electric diad with definite polar moment orientated in definite direction,—and with this the new schemes come into contrast. The following general remarks may be useful towards comparison of these two fundamentally different conceptions. Other reasons will incidentally emerge, besides the results of analysis by X-rays, favouring the view that facts of nature point rather towards the atomic mode of crystal structure.

These considerations deal with fundamental ideas that must be reckoned with, on an ionic view, apart from the detail essential to the science of crystallography. It will however be convenient to fix attention by simple cases; for example, the typical space-lattice for an alkaline halide such as rocksalt NaCl, in which the crystalline unit has been ascertained to be

---

[*] Rather the ion, which may be an atom or a group of atoms. The problem as to how stability of the crystalline structures is assured by interaction of the electric forces with the forces concerned in the packing of the molecules is outside the scope of this note, which is concerned with analysing the results of various types of structure.

for each ion a "face-centred" cube, the corners for one component being the centres of the edges for the other. In the two structural plane diagrams, which represent the atoms projected perpendicular to a diagonal, the letters *a* may represent the negative atomic ions Cl, and the letters *b* the positive ones Na. In the second diagram a diagonal boundary face constituted of Cl atoms is succeeded by parallel planes, alternately of Na and Cl, and finally one of Na, these planes of like atoms being obliquely placed: this diagram invites description at sight as that of a very strongly polarised aggregate, each bipolar element being made up of a Cl combined with an adjacent Na below it. On the other hand in the first diagram the boundary face is composed of both kinds of ion in equal numbers: and if the aggregate is regarded as sorted into bipolar elements, each consisting of an atom in one face and the adjacent atom in the next parallel plane, the polarities of the consecutive elements in all directions are equal and opposite, so that the mass appears as unpolarised when the effective element of volume is large enough to include a considerable number of such bipolar elements.

a	*b*	*a*	*b*
b	*a*	*b*	*a*
a	*b*	*a*	*b*
b	*a*	*b*	*a*

a			
b	*a*		
a	*b*	*a*	
b	*a*	*b*	*a*
	b	*a*	*b*
		b	*a*
			b

These two modes of description of the same lattice can of course be only apparently in contradiction. In the Poisson-Kelvin conception of a polarised medium, the polarity can be replaced, except for locally constitutive purposes,[*] by equivalent distributions of simple density throughout the volume and over the surface. In a crystal, in which the structure is strictly uniform, the volume density is zero on both these modes of polar specification of NaCl: and to reconcile them it is the surface density that must claim attention.

These regular polarities are special, as being far more precise distributions than the fortuitous aggregates of polar molecules that are the foundation of the usual theory of a polarised medium, as constructed by the process of averaging over an element of volume.

By a mathematical device familiar in the elementary analysis of optical diffraction, it is possible to reduce the polarity overtly associated with the second NaCl diagram. We have only to regard each ionic charge as divided into two halves, of which one is associated with half of the next ion in front and the other with half of the next ion in rear along a row, thus forming two adjacent bipolar elements, of opposite signs, and therefore practically cancelling each other except as regards local physical quality of the substance. In this way the electric system is reduced to two surface densities of opposite signs, represented by half the distributions of ions on the opposite boundary planes. The factor one-half at first glance may seem to be a discrepancy between this regular discrete distribution of poles and the Poisson-Kelvin averaged theory of continuous polarisation: the difference may be traced to the circumstance that the polar pairs in the second diagram occupy only half the intervals along a row of atoms, the rest of the length being vacant.

---

[*] Compare for example, 'Æther and Matter,' 1900, p. 252.

The equivalent surface density of ions, here contemplated over limited areas, would involve enormous electric fields in the substance between these intensely electrified boundary planes, such as could not possibly subsist. It must be compensated locally on each element of surface. If conducting quality of metallic type is absent so that free electrons are excluded, this can only be by a distribution of ions of the opposite kind added on outside the surface layer, but only of half the number that would make up a complete stratum of the crystal. We conclude that a boundary plane, in a crystal, such as contains only one kind of ion, must be a ragged surface instead of a smooth one; for it involves the presence of an additional half stratum of the other kind.* When a crystal face grows, the ions must be laid down alternately, and also so that successive pairs constitute opposed polar elements: thus the modes of growth of the two types of face would be different: and the same would apply to modes of dissolution.

The crystalline structure of various metals has been found by W. L. Bragg and other investigators to agree with this specially simple face-centred cubical space-lattice of Na, which in the scheme of NaCl is electrically compensated by the similar but displaced lattice for Cl. In case of the pure metal also, electric compensation must be provided, most naturally perhaps† by locating negative electrons more or less permanently in the positions occupied in the salt by the Cl ions.

Hitherto there is on this view no provision for the liberation of latent surface electric charge by change of temperature or by stress, as is conspicuous in crystals such as tourmaline. For the number of ions or electrons required for a neutralising surface distribution depends only on the number of ions in the crystal, and is not affected by change of its form or dimensions.

---

* An oblique artificial face should exhibit similar features as regards electric excitation, unless it were made up of minute facets with boundary planes in the principal directions of the crystal.

† Lindemann; Born and Landé; also Borelius, 'Phil. Mag.,' 1920.

This additional feature will be introduced if the positive and negative ions are not equidistant along their lattice rows. Thus for the line of ions represented in the diagram, we have to pair into bipoles by resolving each charge, say the one at B, into unequal fractions, so that the moments formed by their association with fractions of the adjacent charges at A and C may balance. The uncompensated charge at the terminal ion A is now e.AB/AC instead of $\frac{1}{2}e$; and similarly at the other end of the chain.

A B   C D   E F
b a   *b a*   *b a*   ...

These residual charges on the surfaces of the crystal must, as before, be masked by local distribution of additional ions or by electrons. But the masking charge no longer stands in the same simple relation to the number of surface atoms: and whenever, by thermal expansion or by stress, the ratio AB/AC becomes altered in the crystal, the compensation at the surface will become deranged, indications of free surface electrification will appear, and persist until they are wiped out by further accretion of compensating charge. An extremely small change in the ratio AB/AC would suffice to cover the facts of pyro- and piezo-electrification.

When each crystalline element is itself polarised with regard to a definite axis, the Poisson-Kelvin theory of a medium polarised *in bulk* can be formulated. But in the configuration here under review it seems to require modification; the positive and negative elements are independent ions, in regulated positions, and it has been seen that we can regard them as associated in pairs in various ways; so that there is no definite unique specification in terms of polarity per unit volume of the medium. But the analytical theory of the charges set free in crystals by stress or change of temperature seems to have hitherto been developed with regard to an assumed specific electric polarity in the crystal, having reference to some definite axis; with the result that cases were outstanding which would not come under the rules,*

---

* It appears that an intrinsic polarity, or regular static collocation of electrons within each molecule, would suffice by interacting with crystalline deformation, to include within the theory the octopolar pyro-electric quality detected long ago in cubic crystals by Haüy: *cf.* Lord Kelvin's 'Baltimore Lectures,' p. 561, or "Æpinus Atomized," 'Phil. Mag.,' 1891.

The formal theory of Voigt ('Krystalloptik,' §§ 128–9) seems to regard only *induced* polarity, which is expressed as a function of deformation *alone*, whether due to stress or to uniform change of temperature, by means of a sufficient number of moduli. It appears from recent experiments by Roentgen, Hayashi, Ackermann, and Lindman, in and had to be referred to internal strain set up by cooling. On the present view such cases perhaps cease to be abnormal; for a direction of polarity cannot be ascribed to the compound lattice. The theory thus opens out into a new form of problem in crystalline tactic; the opposing ions within the crystal have to be paired, either as wholes or in fractions, so that the moments of the adjacent pairs shall everywhere cancel out. Then uncompensated ions on the surface remain over unimpaired, except so far as they are masked by additional electrons or ions brought there from without; they would give rise to the emergence of free surface charge when the physical state of the crystal is changed.

Such release of latent electricity should not appear on surfaces constituted by equal numbers of the two ions. The specification of its distribution, peculiar as being locally conditioned, should of course be definite and unique, by whatever process of pairing it is reached. The electric abnormalities caused by cracks and clefts on a crystal surface would now arise naturally, without requiring local stress due to unequal heating, as had to be suggested.

Dielectric excitation in a uniform electric field would now be represented by the positive and negative component lattices being drawn towards opposite directions by the field, with resulting electric moment depending on their *relative* displacement. For NaCl in a uniform field of electric force the two similar lattices of the Na atoms and the Cl atoms are thus displaced relative to each other without deformation of either though with very slight change of structure; and this applies generally, for all component lattices are similar. Here also the electric displacement could be fixed and compensated, by local surface charge masking the polarisation—for example by sweeping a flame over the surfaces while the inducing electric field is in action. In the alternating electric fields of radiation it is, on final analysis, this oscillatory motion relative to each other* of the component ionic lattices that influences the speed and mode of propagation of the waves.

If the crystal has not sufficient symmetry (in fact, in the lattices that have spiral features) this relative displacement of the two ionic groups in the element of volume, by an applied electric field, may be one of twisting type, involving a rotational part to which is to be ascribed the chiral structural

---

which special precautions were taken as to uniformity of temperature, that these moduli are not much, if at all, affected by temperature (*cf.* Voigt, 'Ann. der Physik,' 1915); as would be natural on the present order of ideas. It was in fact held by J. and P. Curie, the discoverers of piezo-electricity, that pyro-electricity is to be regarded as a result of deformation alone and not as essentially thermal.

* Only partially regular and elastic in the case of metals: the super-conductance of K. Onnes has been ascribed (Lindemann) to complete regularity owing to thermal agitations being suppressed.

rotation of plane of polarisation of light transmitted through the crystal. There are thus the two modes of origin of this chiral optical rotation. One depends on crystalline structure alone, and disappears on fusion of the crystal. In the other the molecule itself, usually a complex aggregate of atoms, undergoes in an electric field this spiral relative deformation of its ionic constituent groups, involving screw quality; and to this exent the optical chirality will remain on fusion or solution, but in altered amount, usually much smaller, as was observed in the case of sucrose by Pocklington. From achiral molecules a crystalline structure of chiral quality can thus be formed. As chiral molecular quality is not of vector type, it could not be annulled by fortuitous orientations as in a fluid, when molecules of opposed chiralities are not present. If the chirality of the molecules is isotropic (*cf.* Lord Kelvin's isotropic helicoid) there would not be much alteration by fusion of that small intrinsic part of the chirality in bulk; if uniaxial, the maximum would be reduced to about one-third.

Double refraction, limited in form* in accord with the dynamical equations of MacCullagh and of Maxwell, will arise from any kind of deviation from cubic symmetry in the natural crystalline ionic lattice, supposed of structure very fine compared with the lengths of the waves. But such defect of symmetry can also be produced by mechanical deformation of a regular crystal: and the double refraction, necessarily slight, thereby introduced (in fact by bending of the lattice) even for regular crystals, must hold a clue to the intimate electric characters of the strain. An atomic lattice must be far more amenable to mechanical stress than a set of molecules could be. In amorphous aggregates, such as glass and gelatine, and even in viscous liquids under forced shearing flow, the deformation by stress of some less regular electric structure is indicated:† in fluid colloids electric ionic phenomena are of course in other ways conspicuous.‡

The magnetic achiral optical rotation of Faraday, common to all matter however regular in its crystalline form, must arise in quite different way. Here the primary phenomenon is change of free vibrational periods of the individual molecules, or perhaps atoms if the molecule is so loose a structure as is here indicated, by the extraneous magnetic field, which is the Zeeman effect: and the optical rotation is an affair of resonance due to the change of these free interacting periods,* after the manner in which ordinary dispersion is produced.†

---

* The Principle of Minimal Action excludes anything more general, except chirality of the types here reviewed, *provided* the waves of radiation are long in comparison with the lattice structure.

† The recent optical study of fused quartz by Lord Rayleigh seems to open out a path in this direction.

‡ For example, the particles in a metallic suspension or (so-called) colloidal solution are now known by X-ray analysis to be crystalline; and a layer of electrons attached and locally compensating the intrinsic polarity would account for the electric charges which are the dominating feature of such suspensions. Over the whole surface this charge would be as much positive as negative, but negative electrons would be attached (in appropriate regions) by preference.

In both types of optical rotation these natural explanations concur with the forms for the equations to which the phenomena are restricted by the principle of Conservation of Energy or rather, more generally, the principle of Minimal Action.‡ Thus if we consider the case of chirally isotropic substances, the alternating electric field (P, Q, R) of the travelling radiation would introduce, through the twisting displacement of the ions, in the molecule or the lattice or both, a concomitant effective magnetisation $k\,d/dt\,(\mathrm{P}, \mathrm{Q}, \mathrm{R})$; so that its magnetic induction (a, b, c) would be given by equations of form $a = \alpha + 4\pi\mathrm{A}$ where $\mathrm{A} = k\,d\mathrm{P}/dt$.

Of the field equations of optical transmission, of type

$$\frac{\partial \mathrm{R}}{\partial y} - \frac{\partial \mathrm{Q}}{\partial z} = -\frac{da}{dt}, \qquad \frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z} = 4\pi\frac{df}{dt}$$

the second could thus be written, on substitution for α, β, γ,

$$\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} = \frac{d}{dt}\left\{4\pi f + 4\pi k\left(\frac{\partial \mathrm{R}}{\partial y} - \frac{\partial \mathrm{Q}}{\partial z}\right)\right\}, \quad \text{where } f = \frac{\mathrm{K}}{4\pi\mathrm{C}^2}\,\mathrm{P},$$

agreeing in form with the equations derived from the Principle of Extremal Action.§ When K is uniform, these equations preserve both the electric and the magnetic inductions free from divergence, as they ought to be.

The general outcome is that a set of component similar lattices of electric ions prescribes in a natural manner the existence and character of the electric properties recognised in crystals: double refraction, pyro- and piezo-electricity, chiral polarisation, artificial double refraction excited by mechanical stress. It would seem even possible to infer, solely from absence of pyro-electric effect in the more symmetrical forms, that in them the crystalline lattice must be formed of ions equably spaced and not combined into molecular groups. For if the molecule were the unit, expansion by heat would alter the distance between adjacent molecules but not, or at any rate not in the same way, the distance between the two electric poles in each molecule: and it would follow by the argument indicated above, that it must alter the amount of the uncompensated surface charge, thus involving pyro-electric phenomena. The only way to avoid this conclusion would be to imagine that the positive and negative ions are so intimately blended in the molecule that it has no resultant moment, after the manner of the first crystal diagram above, until the positive and negative interlacing groups in it are displaced relatively to each other by an imposed electric field.

---

* *Cf.* 'Æther and Matter,' p. 353.

† The static magnetisation of crystals must be an affair of the orientated individual atoms, influenced only slightly by the proximity of the adjacent ones.

‡ 'Æther and Matter,' pp. 197, 206.

§ *Loc. cit.*, § 133. The expression there given for the velocity of propagation is misprinted: *cf. infra* at the end of this paper.

Hysteresis in dielectrics, producing the residual charge of condensers, was referred by Maxwell to heterogeneity operating along with transfer of masking charges by conductance. Prompted by this ascription, Rowland found that in crystal dielectrics there was no residual charge. One can imagine dielectric fragments A, B, ... separated by a medium of different dielectric and also slight conducting quality. In a field of force, A, B, ... are polarised, but the local field is gradually diminished by conductance between A and B conveying masking charges to their surfaces. When the coatings of the condenser are discharged, the whole field throughout the dielectric is suddenly reduced, producing a sudden fall in the polarisation of A, B, ..., but not in the surface charges that masked it. The latter, now uncompensated, gradually make their way back in the reverse direction, involving accumulation of residual charge on the coatings. The actual residual charge would thus come from a small depth, but is permitted to flow back by local compensation in deeper layers. In an alternating field such as that of radiation, conductance would operate by a sort of approximation to a Grotthus chain. But in a very dilute electrolytic solution, the ionic atoms must apparently be wholly isolated from one another, whether the dielectric quality of the solvent be due to incipient atomic lattices or to orientation of polar molecules.

The internal charges in the dielectric being locally masked, the capacity of a condenser as regards sensible charge on its coatings should not be variable, as R. Kohlrausch in fact verified long ago.

The ferromagnetic hysteresis elucidated by Sir A. Ewing, and so vividly illustrated by his working models, is necessarily of different type, for there are no magnetic ions to compensate the polarisation. It requires the molecule, with its own polar magnetic moment, as unit of the structures which acquire instability under increasing extraneous force and finally fall over; yet repeated reversals do not seem to wear down completely such tendency to instabilities, which must be somehow reconstituted.

Apart from hysteresis, it has been shown long ago* that the law of Curie, paramagnetism inversely as absolute temperature, is the expression of general thermodynamic principles without involving any special theory.

---

* 'Phil. Trans.,' 1897, A, p. 287.

On the same lines it has been argued* that, if electric polarisation were an affair of orientation of existing polar moments, $K-1$ should tend to vary inversely as temperature, which it seems to show no sign of doing.

The specific optical rotations of the usual active substances are of the same order of magnitude: and ideas advanced to represent their origin must be expected to throw light on that fact.

The equations expressed above lead, on elimination so that the electric force remains, to the type

$$\nabla^2 P = \frac{\partial^2}{\partial t^2}\left\{\frac{K}{c^2} P + 4\pi k\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\right\}.$$

Thus for radiation travelling along z, in which therefore both the circuital vectors are on the wavefront xy,

$$\frac{\partial^2 P}{\partial z^2} = \frac{K}{c^2}\frac{\partial^2 P}{\partial t^2} - 4\pi k\frac{\partial^3 Q}{\partial z \partial t^2},$$

$$\frac{\partial^2 Q}{\partial z^2} = \frac{K}{c^2}\frac{\partial^2 Q}{\partial t^2} + 4\pi k\frac{\partial^3 P}{\partial z \partial t^2},$$

so that if U represents $P+\iota Q$, $\iota = \sqrt{-1}$,

$$\frac{\partial^2 U}{\partial z^2} = \frac{\partial^2}{\partial t^2}\left(\frac{K}{c^2} + 4\pi k\iota\frac{\partial}{\partial z}\right) U,$$

indicating waves of the two circular types. If they are harmonic wave-trains involving the exponential factor

$$\epsilon^{\iota pt - \iota mz}, \qquad m = 2\pi/\lambda = p/c',$$

then

$$m^2 = \frac{p^2}{c^2}(K + 4\pi k c^2 m),$$

giving for the velocities of propagation p/m of the two circular radiations of the same period $2\pi/p$ the approximate values

$$c' = cK^{-\frac{1}{2}}(1 \pm 2\pi k K^{-\frac{1}{2}} cp).$$

The optical rotation for unit depth of the medium is the difference in π/λ or in $\frac{1}{2}p/c'$ for these two radiations, and this is $\pi p^2 k$.

Now approximating very coarsely, if there are N molecules or chiral elements per unit volume, in each of which positive and negative ions (say single charges e) are drawn asunder a distance δl_0 by the electric field and screwed through an angle θ_0, then omitting the common factor $\epsilon^{\iota pt}$

$$k\frac{\partial P}{\partial t} = \tfrac{1}{2}\iota N e r^2 \theta_0 p, \qquad \frac{K-1}{4\pi c^2} P = N e\,\delta l_0,$$

where r represents a mean radius of transverse section of the molecule (or of a spiral filament of the lattice) so that the magnetic moment is say $\frac{1}{2} e r^2\, d\theta/dt$ for each chiral unit.

---

* *Loc. cit.*, p. 283.

On division N disappears and

$$\frac{K-1}{4\pi c^2 k} = 2\frac{\delta l_0}{r^2\theta_0},$$

so that the optical rotation per unit depth is

$$p^2\frac{K-1}{8c^2}\frac{r^2\theta_0}{\delta l_0}, \quad \text{or} \quad \frac{\pi^2}{2\lambda^2}(K-1)\,r\,\frac{r\theta_0}{\delta l_0}.$$

If r is of the order of the diameter of a molecule, say 10^{-8}, this reduces to the order of $10r\theta_0/\delta l_0$ for visible light.

Actually the optical rotation is of the order of two right angles per centimetre for quartz, and seems to range to several times less for pure active liquids. This indicates that the angle of pitch of the relative screw displacement of the positive and negative components of the structures should range around $\frac{1}{10}$, so that the chiral element may twist under an electric field about one-tenth as much as it elongates. This indication is extremely rough; but bearing in mind that powers of the velocity of radiation are involved in the estimate, one may claim that it does not detract from, but rather substantially supports, the type of picture of the phenomena which has been under review.

The Colour of the Light from the Night Sky.

By Lord Rayleigh, F.R.S.

(Received February 7, 1921.)

The general light of the sky at night is very faint, and beyond the observations of Campbell[*] and Slipher,[†] who could always detect the aurora spectrum line in temperate latitudes, nothing appears to be known of its spectroscopic or chromatic character. Slipher, indeed, suggested that the variability in its brightness recorded by Yntema might be accounted for by attributing the light mainly to a variable aurora. But this view has been controverted, rightly as I think, by Fabry.[‡] My own observations, now to be recorded, indicate that at the time and place where I was working little, if any, of the light could be attributed to the green aurora line.

The great difficulty of spectroscopic observations is, of course, the want of light; though if the light were monochromatic, as, for instance, if it were all concentrated in the aurora line, this difficulty would be much mitigated. I thought it worth while, however, to try if anything could be learnt by a cruder method, which possesses some advantages even if the light is not very feeble.

---

[*] 'Astr. Soc. Pacific.,' vol. 29, p. 218 (1917).

[†] 'Astr. Jour.,' vol. 49, p. 266 (1919).

[‡] 'Astr. Jour.,' vol. 50, p. 308 (1919).

A series of small coloured filters were arranged so as each to transmit a limited portion of the spectrum. These were placed in a suitable frame, which held in addition a neutral-tinted photometric wedge of carbon in gelatine, made by Mr. Sanger-Shepherd. This wedge was graduated in equal parts, according to Hunter and Driffield's method. (Scale-reading 10, light transmitted = 1/10; scale reading 20, light transmitted = 1/100.) A photographic plate was placed in the frame, so that part of it was under the various colour filters and another part under the wedge. This could be exposed to the night sky or other source of light.

After development, the photographic intensity under each coloured screen will match the intensity under some point on the wedge. The plate, therefore, gives a record of what fraction of the total photographic effect due to the source is transmitted by each screen. For comparable results with different sources of light, it is of course necessary to keep to the same kind of photographic plate.

The great advantage of the method is that it allows light to be used from a cone of very large angular aperture. With a spectroscope we cannot, under the most favourable circumstance, use an aperture ratio of more than 1 : 2. But a plate under the coloured filters may be exposed to the entire hemisphere if desired. Moreover, when faint impressions are concerned, a large patch of defined shape and uniform, even if small, intensity practically gives much more confidence to the experimenter than a small and faint spectrum. The information given is of course of a rather different kind.

Coming now to details. The coloured screens were :—

	Transmission.
1. Ultra-violet glass, supplied by Messrs. Chance	3300–3800
2. Cobalt blue glass. Several thicknesses	4000–4500
3. Bluish-green glass	4300–5300
4. Green glass	5100–5900
5. Flavazine	5500–7000
6. Flavazine, with didymium glass	5500–5630 and 5930–7000
7. Wratten's red filter, No. 23A	5800–7000

These were chosen, so far as possible, to give nearly the same photographic density with sunlight. This is desirable, because the range of intensities which the plate can deal with is not unlimited. The green and red filters, which transmit portions of the spectrum to which the plate is comparatively insensitive, were chosen to give as large transmission as was consistent with isolating a fairly narrow range of the spectrum.

Filter No. 6 was designed to isolate as nearly as possible the green aurora line, if present. No filter could be found to cut out the red without at the same time involving serious loss of green.

Ilford special rapid panchromatic plates were used. These were far more effective in the red, yellow, and green regions than any other plates I have tried. All the results recorded were obtained with a consignment of plates bearing the same number, and therefore, presumably, as nearly alike as may be.

An ink mark on the under face of the wedge gave a fiducial mark on the negative, and starting from this, the wedge scale was copied by hand on to the developed film, marking it on with a fine pen.

The photometric work on the developed plates was done with the photo-electric apparatus described in connection with the scattering of light by gases.* Briefly, it depends on a photo-electric cell with battery and galvanometer. The plate rests on a table, in which is a slit. Above is a bright light focussed on the slit with a lens. The rays diverging beyond the focus fall on the photo-electric cell. The deflection of the galvanometer depends on the opacity of the photographic film. The impression through a colour filter was placed on the slit, film downwards, and the deflection noted. The impression of the wedge was then substituted, with the length of the slit at right angles to the length of the wedge. The plate was then slid along until the previous deflection was recovered, and the position of the slit on the wedge scale noted.

On a given negative, the result was usually recovered to within about 0·2 of a scale division, but, with another negative, discrepancies as large as 1 scale division (representing a factor of 1·26), or even more, were observed. These are due either to inequalities in the photographic film† or to changes in the quality of light received from the source, due to varying atmospheric conditions.

The photo-electric measurement of a given negative was very satisfactory, and far preferable to visual photometry.‡

The first series of experiments was made with the plate exposed to the entire sky, so far as it was not blocked out by trees and buildings. This method is open to the objection that some of the rays pass very obliquely through the coloured screens, with the result that the effective thickness of the latter becomes somewhat indefinite. This, however, does not seem to vitiate the result appreciably, as subsequent experience showed; while an important advantage is gained in shortening the exposure.

---

* 'Roy. Soc. Proc.,' A, vol. 97, p. 440 (1920).

† See the remarks in my paper, *loc. cit.*, p. 442. The plates used in the present work were coated on common glass, not plate glass.

‡ The latter can, no doubt, be applied to smaller areas on the negative, which is important in some investigations.

Comparison of the night sky with the clear twilight sky (Table I) shows clearly how much poorer the latter is in yellow and red light. The wedge readings for the two yellow and the red filter differ by about 4 scale divisions, showing that, for the night sky, the yellow and red components account for 2·5 times as large a percentage of the whole photographic effect as in the twilight sky.

Table I.—The Figures given are the Wedge Readings, obtained as described.

	Starlight sky.		Twilight, clear sky. August 17.
	August 9.	September 6–7.	
Ultra-violet glass	10·6	10·0	9·9
Cobalt glass	13·1	13·3	12·4
Bluish-green	12·5	12·2	11·8
Green	14·7	13·7	16·1
Flavazine	8·5	8·0	18·1
Flavazine with didymium	12·1	10·9	16·1
Red	10 7	10·0	15·0

The next series of experiments were designed to include in the comparison the direct light of the sun and moon. These give approximately parallel rays, which traverse the absorbing screens normally. For comparison with them, it is desirable to restrict the cone of rays from the night (or day) sky as far as other conditions will allow. An aperture, 15 inches in diameter, was placed 15 inches from the plate. This aperture was the smallest that could well be used. The work was carried out in August and September and early October, when the hours of complete darkness are very limited. Summer time imposes a further obstacle. The moon must be below the horizon and the night should be clear. With the aperture ratio 1 : 1, as described, about nine hours' exposure were needed to give satisfactory intensity.

Two plates with this exposure were all I could get before my enforced return to London in October. The exposures were made at Beaufront Castle, 3 miles from Hexham, Northumberland, and also at Terling Place, Essex. There was no appreciable interference from town light, diffused by the sky.

In making exposures to direct sunlight, the difficulty was to reduce the photographic effect sufficiently. The ideal method would be to use a pinhole camera, giving an image of the sun which would cover the quarter plate used, but this involves an impractically long camera. To work at a reasonable distance, more divergence of the rays is essential. It was got by means of a quartz lens, of 7 mm. diameter and 9 mm. focus. The plate was placed on the axis of the lens, at about 60 cm. distance, when the divergent cone of rays from the image covered an area more than large enough. The lens was mounted in the end of a blackened box, with a kodak photographic shutter over it. The plate holder, carrying the coloured screens and photometric wedge, was introduced at the other end of the box. A simple shadow arrangement was provided for pointing the axis of the lens at the sun.

A quartz lens was used, so as to avoid any absorption of ultra-violet light, or any risk of introducing colour which might disturb the comparison of sunlight with other sources.

Exposures were also made on direct moonlight. For these no special arrangements were needed, as an exposure of several seconds proved suitable.

The results are given in Table II.

Table II.

	Starlight sky.			Blue sky at sunset.		Blue sky, Sept. 24, 1 p.m.
	Sept. 15–16.	Oct. 8, 9, 10.	Mean.	Sept. 27.	Sept. 28.	
Ultra-violet glass	13·2	11·9	12·6	12·3	15·4	10·4
Cobalt glass	14·9	14·8	14·8	13·5	12·8	14·0
Bluish-green	11·8	12·3	12·0	11·9	13·3	12·8
Green	13·5	14·7	14·1	16·3	17·2	17·6
Flavazine	8·5	8·5	8·5	12·9	15·1	13·9
Flavazine with didymium	11·5	11·6	11·0	16·6	17·0	18·0
Red	9·9	11·9	10·9	14·3	16·2	16·3

	Sun, 33° up.			Moon, Sept. 26, 33° up.		Basalt, fresh.	Basalt, weathered.
	Sept. 22.	Sept. 28.	Sept. 28.				
Ultra-violet glass	14·9	15·8	14·9	14·9	15·0	14·9	16·2
Cobalt glass	14·3	15·2	14·6	15·2	15·0	14·4	15·6
Bluish-green	11·7	12·9	13·2	11·4	12·0	11·3	11·5
Green	13·2	14·1	15·5	14·0	14·2	13·1	12·3
Flavazine	7·0	8·0	8·8	7·5	7·4	7·8	5·5
Flavazine with didymium	10·7	11·0	12·2	11·3	10·7	10·6	8·3
Red	9·0	10·8	11·5	9·5	8·7	9·3	7·4

In considering these results, it will be noticed in the first place that the wedge readings for cobalt glass and bluish-green glass respectively are nearly the same whatever the source of light. This merely expresses the fact that by far the greater part of the photographic effect of the spectrum is due to the kind of light which these glasses can transmit. In any case the ultra-violet or red ends of the spectrum produce a comparatively small fraction of the whole effect, consequently the absence or presence of these constituents hardly affects the wedge reading for the blue.

It is in the case of the ultra-violet at one end of the spectrum and the yellow and red at the other that enough variability is to be expected to give information as to the constitution of the source.

The ultra-violet constituent tends to be variable for plates obtained on different occasions. This is not surprising when we consider that this constituent is largely affected by atmospheric absorption, and is therefore sensitive to atmospheric conditions, even when the sky is apparently clear. For the same reason this constituent will be largely affected by the altitude of the source. Thus the sun and moon, both exposed to at 33° altitude, are at a disadvantage compared with the zenith sky.

For these reasons not much can be inferred from the ultra-violet data. It is noteworthy, however, that the zenith sky seems much richer in this constituent when the sun is high, than at sunset. This may be accounted for by atmospheric absorption of the rays from the low sun before they reach the observer's zenith, and are scattered towards him by the air.

The great superiority of the night sky over the day sky in yellow and red light is again apparent, amounting as before to about 4 scale divisions (a factor of 2·5).

The night sky also differs very little in quality from direct sunlight. It is only in the ultra-violet that any definite distinction is apparent, and as already remarked, the zenith sky would have an advantage in this region on account of less atmospheric absorption.

No definite distinction can be made from these experiments between the chromatic constitution of sunlight and that of moonlight. The numbers found are the same within the limits of experimental error.

We may say therefore that the night sky is found to be of the same colour as direct sunlight or moonlight, but much yellower than the clear day sky.

Colour of Moonlight.

What we can see of the moon's surface suggests that it consists of volcanic rock, and the light is, of course, diffusely reflected sunlight. It occurred to me, in connection with the present experiments, to include in the comparison sunlight diffused from specimens of terrestrial volcanic rock. These exposures were made when the sun was 33° up, so that atmospheric absorption should be as nearly as possible the same as when the moonlight exposures were made. The photographic plate was placed in the box used for exposures to direct sunlight, but the quartz lens was removed. The iris diaphragm attached to the shutter was adjusted so as to subtend half a degree at the photographic plate—the same angular diameter as the moon—and it was backed with the specimen of rock in full sunlight placed near the aperture. The same exposure was given as was given to direct moonlight. Only two exposures were made, one to a specimen of fresh basalt, and another to a weathered ochreous surface of the same. The light from the fresh basalt cannot be distinguished definitely in quality from moonlight or sunlight, but the light from the weathered basalt appears to be distinctly yellower than moonlight. The experiments are too few for a final conclusion, but they suggest that the surface of the moon is more like the fresh than the weathered specimen. This is in accordance with the absence of a lunar atmosphere.*

Visual Observations on the Night Sky.

The conclusions reached, photographically, on the night sky are entirely confirmed by visual observation.

The photometric comparison of lights of different colours is, under ordinary circumstances, difficult and embarrassing, and little definite significance attaches to the results. But with lights below a certain intensity this difficulty disappears, since there is no longer the power of colour discrimination, and vision is monochromatic. A confident decision can then be given as to which of two lights appears brighter, whether they are of the same spectral quality or not. This is, of course, well known, and is fully elaborated in works on colour vision.

Two gelatine films on glass were prepared, one dyed yellow with flavazine, and the other with methylene blue, the intensity of the latter being adjusted by washing out some of the colour until right for getting the effects which will now be described. The two films were mounted edge to edge at the end of a pasteboard tube, so that, when the tube was directed to the sky, the circular field was seen divided into two along a diameter.

During the daytime the yellow film was considered by all observers to be the brighter. As twilight advanced, the Purkinje phenomonon came into evidence, and the blue film became brighter. This remained the case when the light had waned so far that the colour sensation had disappeared. As the stars came out, the predominance became less marked, and, before the Milky Way was distinguished, there was equality. Finally, when the Milky Way was conspicuous, the yellow film was notably brighter, whether the tube was pointed to the Milky Way or to other parts of the sky.

---

* These experiments on rock were hampered by the fact that the photometric apparatus was not at hand when the exposures were made. They are merely of a preliminary character.

Another pair of films was prepared, giving equal intensities when the night sky was viewed. In this case, still more than in the previous one, the blue half appeared brighter while any twilight remained.*

It was desired to confirm visually the conclusion reached by photography that the colour of the night sky was nearly the same as that of direct sunlight or moonlight. Moonlight was chosen, as being of more convenient intensity. It was found that the yellow and blue films, which appeared equally bright against the night sky, also appeared equally bright when viewed against a white cloth exposed to direct moonlight. This background was, however, somewhat too bright, and, in order to get rid of colour perception, it was found advisable to reduce it twelve-fold by means of Fox Talbot's revolving sector. It was then impossible to distinguish between the two coloured films, which appeared of exactly equal intensity.

The night sky was examined on many occasions through the pair of films, which gave equal brightness. I never observed any marked deviation from equality, at considerable altitudes, though low down the yellow often had an advantage. This is probably attributable to selective atmospheric absorption. The experiments gave no evidence of a variable colour such as might be produced by an aurora of variable intensity.

I may add that my own attempts to see or photograph the aurora line on ordinary nights have had no success. All that I could see was a continuous spectrum.

[*Note added, March* 4.—I have since succeeded in regularly photographing the aurora line.]

Summary and Conclusion.

To get information as to the chromatic character of the light from the sky at night, photographic exposures were made under coloured media selected for isolating various parts of the spectrum. By comparing the photographic densities obtained, it was concluded that the night sky was much yellower or less blue than the (clear) day sky. Comparison with direct sunlight or moonlight showed that the night sky was of the same quality as these.

Visual comparisons through coloured films showed that a blue film, which was equally bright with a yellow one against the night sky, was brighter against the twilight sky. These comparisons were not embarrassed by colour differences, because the light was so faint as to give purely monochromatic vision. The two films matched one another equally well, whether they were seen against the Milky Way or against other parts of the sky. In the photographic work no special attention was paid to this point, the exposures being made to a considerable area of sky around the zenith. The diurnal motion, of course, brought successive areas of the sky into action during the long exposures.

---

* A preliminary notice of this result was published in a letter to 'Nature,' dated August 20, 1920.

One theory of the light of the night sky attributes it to sunlight scattered by a very rare gaseous atmosphere, situated so high up as to be outside the earth's shadow. The present observations are contradictory of this theory, which would require the night sky to have the same colour as the day sky.

The comparative absence of polarisation* leads to the same conclusion.

The requirements, as regards colour and polarisation of the light, would be satisfied if we regarded it as coming from an unresolved background of stars. They would equally be satisfied if we regarded it as due to sunlight scattered by meteoric matter. The conclusions of van Rhyn† favour the latter alternative. He regards the general light of the sky as an extension of the zodiacal light.

* See Rayleigh, 'Astr. Jour.,' vol. 50, p. 227 (1919); also Babcock, *ibid.*, p. 228.

† 'Astro. Phys. Jour.,' vol. 50, p. 356 (1919).

Linear Transformations and Functions of Positive Type.

By J. Mercer, M.A., D.Sc., Fellow and Lecturer of Christ's
College, Cambridge.

(Communicated by Prof. E. W. Hobson, F.R.S. Received December 10, 1920.)

In a communication received by the Society on April 12 last[*] I gave in outline a proof that a symmetrisable function $\kappa(s,t)$, which belongs to one of certain classes, admits an expansion in terms of its bi-orthogonal fundamental functions. Two classes were explicitly mentioned, viz., (i) the functions defined by $\kappa(s,t) = \alpha(s)\gamma(s,t)$, and (ii) those defined by

$$\kappa(s,t) = \int_a^b \alpha(s,x)\gamma(x,t)\,dx.$$

It was assumed in each case that $\gamma(s,t)$ was of positive type in the square $a \leq s \leq b$, $a \leq t \leq b$; in the first case, that $\alpha(s)$ was a function defined in the interval (a,b); and, in the second case, that $\alpha(s,t)$ was a symmetric function defined in the square $a \leq s \leq b$, $a \leq t \leq b$. When $\kappa(s,t)$ belongs to either of the classes mentioned and certain conditions of continuity[†] are satisfied it was shown that

$$\kappa(s,t) = \sum_{n=1} \frac{\phi_n(s)\psi_n(t)}{\lambda_n} + \sum_{n=1} \frac{\xi_n(s)\eta_n(t)}{\upsilon_n}, \tag{1}$$

where (i)

$$\phi_1(s),\ \phi_2(s),\ \ldots,\ \phi_n(s),\ \ldots$$
$$\psi_1(s),\ \psi_2(s),\ \ldots,\ \psi_n(s),\ \ldots$$

is a bi-orthogonal system of fundamental functions for the interval (a,b), and $\phi_n(s)$, $\psi_n(t)$ satisfy the homogeneous equations

$$\left.\begin{aligned} \phi_n(s) &= \lambda_n \int_a^b \kappa(s,t)\,\phi_n(t)\,dt, \\ \psi_n(t) &= \lambda_n \int_a^b \psi_n(s)\,\kappa(s,t)\,ds\,; \end{aligned}\right\} \tag{2}$$

(ii) the functions $\eta_n(s)$ $(n = 1, 2, \ldots)$ constitute an orthogonal system of functions for the interval (a,b); and (iii) each of the pairs of functions $\{\phi_m(s), \eta_n(s)\}$, $\{\xi_m(s), \psi_n(s)\}$, $\{\xi_m(s), \eta_n(s)\}$ is orthogonal for the interval (a,b), whatever be the values of m and n. The symbols υ_n $(n = 1, 2, \ldots)$ denote certain positive constants which, by suitable modification of the functions $\xi_n(s)$, may all be taken equal to unity.

---

[*] See 'Proceedings,' A, vol. 97, pp. 401–413.

[†] For $\gamma(s,t)$ the condition was that the function should be continuous with respect to s for each fixed value of t in (a,b).

It will be evident that a function which belongs to either of the above-mentioned classes is obtained from $\gamma(s, t)$ by a particular kind of linear transformation, *the transformation being with respect to the variable s*. Again, a symmetrisable function such as $\alpha(s)\gamma(s,t)\beta(t)$ is obtained by subjecting $\gamma(s, t)$ to two linear transformations, namely, one with respect to each variable. We are thus led to consider a general class of functions, each of which is obtained by subjecting a function of positive type to linear transformations; and to enquire under what circumstances a function of this class admits expansion in the form (1). The present communication has its origin in an attempt to answer the enquiry just stated.

In the first section a summary of the theory of linear transformations due to F. Riesz* is given, and further developments required for the purpose in view are given. At the commencement of the second section functions of positive type are defined, which are more general than those hitherto considered. The theory of quadratic forms with an infinite number of variables is then applied to show that, when $\kappa(s, t)$ is obtained by subjecting one of the functions of positive type to a symmetric couple of linear transformations, (1) is valid almost everywhere,† provided further series involving null functions are added on the right. In general, the equations (2) are not satisfied at all points of (a, b), but they are always satisfied *almost everywhere*. It is further shown that, when the function of positive type has the continuity property assumed in the previous communication, and the transformations are of a restricted character, the expansion of $\kappa(s,t)$ is valid everywhere in the square $a \leq s \leq b$, $a \leq t \leq b$. Finally, it appears that, when the transformation with respect to t is still further restricted, the expansion of $\kappa(s, t)$ is of the form (1).

Linear Functional Transformations.

§ 1. Previous to giving the proof of the expansion theorems, it will be convenient to summarise briefly certain definitions and results which belong to the theory of linear functional transformations.

Let F denote the class of real functions which have their squares summable in the interval (a, b). It is to be observed that this hypothesis does not imply that a function of F has necessarily a single finite value at each point of (a, b). A function of F may be undefined, or multiple valued, at points of the interval which form a set of zero measure; further, the function may be regarded as having the values $+\infty$ or $-\infty$ at points of a similar set.‡ In what follows a function will be said to be *well-defined* in (a, b) if it has a single finite value at each point of the interval.

---

* 'Untersuchungen über Systeme integrierbarer Funktionen Math. Annalen,' vol. 69, pp. 449–497. See particularly § 12. The case $p = 2$ only is considered here.

† This statement is correct, but it will be found that the result obtained is of a slightly more general character.

‡ *Cf.* Riesz, *loc. cit.*, p. 454.

Let Ω be a transformation which converts each member of F into a member of the same class, this member being unique save possibly for an additive null function. If we denote the independent variable by s, and a member of F by $f(s)$, a function which results when the transformation Ω is applied to $f(s)$ is denoted by $\Omega[f(s)]$, or by $\Omega_s[f(s)]$, when it is necessary to indicate the variable with respect to which the transformation is carried out.

A transformation Ω is said to be distributive if, for any two functions $f_1(s), f_2(s)$, which belong to F, and for any two constants, c_1, c_2, we have

$$\Omega[c_1 f_1(s) + c_2 f_2(s)] = c_1 \Omega[f_1(s)] + c_2 \Omega[f_2(s)], \tag{3}$$

almost everywhere in (a, b). The transformation is said to be *bounded* if, for any function, $f(s)$, which belongs to F,

$$\int_a^b \{\Omega[f(s)]\}^2 ds \leq M^2 \int_a^b \{f(s)\}^2 ds,$$

where M is independent of $f(s)$. The lower bound of the values of M, for which the inequality just written is true, may be called the *bound* of the transformation. A transformation which is both distributive and bounded is said to be *linear*.*

As examples of linear transformations we have

$$\text{(i)} \quad \Omega[f(s)] = \alpha(s) f(s),$$

where $\alpha(s)$ is bounded and summable in (a, b), and

$$\text{(ii)} \quad \Omega[f(s)] = \int_a^b \alpha(s, x) f(x)\, dx,$$

where $\alpha(s, t)$ is a function which has its square summable in the region $a \leq s \leq b$, $a \leq t \leq b$. It will be observed that, in each of the examples cited, the function $\Omega[f(s)]$ may not be defined at points of a set of zero measure. The function, however, belongs to F, as was explained above.

§ 2. A transformation, $\bar{\Omega}$, which is related to Ω in such a way that

$$\int_a^b \Omega[f_1(s)] \,.\, f_2(s)\, ds = \int_a^b f_1(s) \,.\, \bar{\Omega}[f_2(s)]\, ds,$$

for any two functions, $f_1(s), f_2(s)$, which belong to F, is said to be *ad-joint* to Ω. It can be shown that any linear transformation has an adjoint, and that the latter is unique, save for an arbitrary null function.† A transformation, Ω, which is adjoint to itself, is said to be *self-adjoint* or *symmetric*.‡

---

* Riesz, *loc. cit.*, p. 477.

† Riesz, *loc. cit.*, p. 478.

‡ Riesz, *loc. cit.*, p. 484. In the examples cited above (i) is always symmetric, (ii) is symmetric when $\alpha(s, t)$ is a symmetric function.

In what follows, we shall be concerned with pairs, Ω, Ω', of transformations, which are such that

$$\int_a^b \Omega\left[f_1(s)\right] \Omega'\left[f_2(s)\right] ds = \int_a^b \Omega\left[f_2(s)\right] . \Omega'\left[f_1(s)\right] ds,$$

for every pair of functions, $f_1(s)$, $f_2(s)$, which belong to F. Two transformations related in this manner will be referred to as a *symmetric-couple*.

It will be observed that, if E is the unit transformation* and Ω is symmetric, then E and Ω constitute a symmetric couple. More generally, any two symmetric transformations which are permutable form a symmetric couple: an important instance of this is the couple defined by $\Omega\left[f(s)\right] = \alpha(s) f(s)$; $\Omega'\left[f(s)\right] = \beta(s) f(s)$.† Other examples are easily constructed.

§ 3. Let $f_n(s)$ $(n = 1, 2, \ldots)$ be a sequence of functions of class F. If a function $f(s)$ of class F can be found such that

$$\operatorname*{Lim}_{n \to \infty} \int_a^b \{f(s) - f_n(s)\}^2 ds = 0,$$

the sequence is said to *converge strongly* to $f(s)$.‡ The function, $f(s)$, when it exists, is unique, save for an additive null function.

It is easy to prove that, if $f_n(s)$ $(n = 1, 2, \ldots)$, $g_n(s)$ $(n = 1, 2, \ldots)$ are sequences which converge strongly to $f(s)$, $g(s)$ respectively, then

$$\operatorname*{Lim}_{n \to \infty} \int_a^b f_n(s) g_n(s)\, ds = \int_a^b f(s) g(s)\, ds.$$

From the bounded property, it follows that, if Ω is any linear transformation, the strong convergence of a sequence, $f_n(s)$ $(n = 1, 2, \ldots)$ to $f(s)$, implies the strong convergence of $\Omega\left[f_n(s)\right]$ $(n = 1, 2, \ldots)$ to $\Omega\left[f(s)\right]$.§ The two results stated allow it to be proved that, when Ω and Ω' are any two linear transformations and $f_n(s)$ $(n = 1, 2, \ldots)$ converges strongly to $f(s)$,

$$\operatorname*{Lim}_{n \to \infty} \int_a^b \Omega\left[f_n(s)\right] . \Omega'\left[f_n(s)\right] ds = \int_a^b \Omega\left[f(s)\right] . \Omega'\left[f(s)\right] ds.$$

§ 4. Let $f_n(s)$ $(n = 1, 2, \ldots)$ be a countable set of functions of class F such that the series

$$\int_a^b \{f_1(s)\}^2 ds + \int_a^b \{f_2(s)\}^2 ds + \ldots + \int_a^b \{f_n(s)\}^2 ds + \ldots, \tag{4}$$

---

* This transformation is defined by $\mathrm{E}\left[f(s)\right] = f(s)$.

† For conditions of summability, see § 1.

‡ Riesz, *loc. cit.*, p. 464; also named 'convergence en moyenné' by E. Fischer.

§ Riesz, *loc. cit.*, p. 464.

is convergent. It follows from a known theorem* that the series

$$\{f_1(s)\}^2+\{f_2(s)\}^2+\ldots+\{f_n(s)\}^2+, \tag{5}$$

converges almost everywhere in (a, b); that $F(s)$, the sum-function of (5), is summable in (a, b); and that $\int_a^b F(s)\,ds$ is the sum of the series (4).

Let $(x_1, x_2, \ldots\ldots x_n, \ldots)$ or briefly (x), be the co-ordinates of a real point in Hilbertian space, and consider

$$f[(x), s] = x_1f_1(s)+x_2f_2(s)+\ldots+x_nf_n(s)+\ldots. \tag{6}$$

For any value of s in (a, b) such that $F(s), f_1(s), f_2(s), \ldots, f_n(s), \ldots$ have all finite values, the right-hand member of (6) is a linear form in Hilbert's sense, (x) being regarded as variable. It follows from what has been said that the member is a linear form for almost all values of s in (a, b). Again, we have

$$\{f[(x), s]\}^2 \leq F(s) \sum_{n=1} x_n^2.$$

almost everywhere in (a, b); hence it is clear that, if (x) is fixed, $f((x), s)$ belongs to the class F. In view of the two properties stated, we shall refer to

$$x_1f_1(s)+x_2f_2(s)+\ldots+x_nf_n(s)+\ldots, \tag{7}$$

as *a linear function associated with the class F*, whenever the series (4) is convergent.

We shall say that the linear function (7) associated with the class F is *well-defined* if (i) the functions $f_n(s)$ are all well-defined and (ii) the series (5) converges at *all* points of (a, b). A well-defined linear function is clearly a linear form in Hilbert's sense for each value of s in (a, b).

§ 5. In virtue of the bounded property of a linear transformation we see that, for any transformation Ω,

$$\int_a^b\{\Omega[f_1(s)]\}^2\,ds+\int_a^b\{\Omega[f_2(s)]\}^2\,ds+\ldots+\int_a^b\{\Omega[f_n(s)]\}^2\,ds \leq M^2\int_a^b F(s)\cdot ds,$$

where M is the bound of the transformation. From this follows the important result that, *if* (7) *is any linear function associated with* F, *and* Ω *is any linear transformation, then*

$$x_1\Omega[f_1(s)]+x_2\Omega[f_2(s)]+\ldots+x_n\Omega[f_n(s)]+\ldots, \tag{8}$$

is a linear function associated with F.

---

* A simple proof has been given by Hobson, 'Proc. Lond. Math. Soc.,' Series 2, vol. 8, pp. 27–29. It is easily seen to be unessential to the proof that the functions $s_n(x)$ should be bounded; further, the validity of the theorem is not affected by the hypothesis that each function $s_n(x)$ is either undefined, multiple valued, or infinite at points of a set of measure zero.

Let us denote the sum of the first n terms on the right of (6) by $f_n((x), s)$. Since

$$\int_a^b \{f((x), s)-f_n((x), s)\}^2\, ds = \sum_{r=n+1} x_r^2 \cdot \int_a^b \mathrm{F}(s)\, ds,$$

it is clear that, when (x) is fixed and n tends to ∞, $f_n((x)s)$ converges strongly to $f((x), s)$. Hence, as n tends to ∞, $\Omega[f_n((x), s)]$ converges strongly to $\Omega[f((x), s)]$. But, since (8) is a linear function associated with F, $\Omega[f_n((x), s)]$ converges strongly to $\sum_{n=1} x_n \Omega[f_n(s)]$. *We thus see that*

$$\Omega[f((x), s)] = \sum_{n=1} x_n \Omega[f_n(s)]$$

almost everywhere in (a, b) at each *point* (x) *of Hilbertian space*. Further it will be seen that

$$x_1 \int_a^b \Omega[f_1(s)]\, ds + x_2 \int_a^b \Omega[f_2(s)]\, ds + \ldots + x_n \int_a^b \Omega[f_n(s)]\, ds + \ldots,$$

is a linear form equal to $\int_a^b \Omega[f((x), s)]\, ds$.

In passing, we observe that, since $f_n((x), s)$ converges strongly to $f((x), s)$ it follows from the remark at the end of § 3 that

$$\operatorname*{Lim}_{n \to \infty} \int_a^b \Omega[f_n((x), s)] \cdot \Omega'[f_n((x), s)]\, ds = \int_a^b \Omega[f((x), s)] \cdot \Omega'[f((x), s)]\, ds,$$

where Ω and Ω' are any two linear transformations.

§ 6. We shall say that a linear transformation is *well defined* if, for each well-defined function $f(s)$ of F, the function $\Omega[f(s)]$ is well defined.

Let us suppose that Ω is well-defined, and that the equation (3) holds exactly at *all* points of (a, b). Then, if $f_1(s), f_2(s), \ldots f_n(s)$ are well-defined functions in (a, b) and $x_1, x_2, \ldots x_n$ are the co-ordinates of a variable point in n dimensional space, we have

$$\Omega[x_1 f_1(s) + x_2 f_2(s) + \ldots + x_n f_n(s)]$$
$$= x_1 \Omega[f_1(s)] + x_2 \Omega[f_2(s)] + \ldots + x_n \Omega[f_n(s)],$$

for *all* values of s in the interval (a, b). This property may be described in the statement that Ω is *precisely distributive in n dimensional space*.

We shall say that a well-defined transformation Ω is *precisely distributive in Hilbertian space* if, corresponding to any well-defined linear function $\sum_{n=1} x_n f_n(s)$, the linear function $\sum_{n=1} x_n \Omega[f_n(s)]$ is well defined and equal to $\Omega[f((x), s)]$, for all points (x) of Hilbertian space and for *all* values of s in (a, b). It will be seen that, *if* Ω *is precisely distributive in Hilbertian space, and if* $f_n(s)$ $(n = 1, 2, \ldots)$ *are well-defined functions such that* $\sum_{n=1} \{f_n(s)\}^2$ *is* *convergent at all points of* (a, b), *then the series* $\sum_{n=1} \{\Omega[f_n(s)]\}^2$ *is convergent at all points of* (a, b).

Let (a', b') be an interval which is part of, or coincident with, (a, b): and let each end point of (a', b') be regarded as belonging to the interval or not, as may be convenient. Let Ω be a linear transformation which is precisely distributive in Hilbertian space. We shall say that Ω is *perfect* in (a', b') if, corresponding to each function $f(s)$ which is continuous in (a', b'), the function $\Omega[f(s)]$ is continuous in (a', b').

As an example of the transformations defined, let us consider

$$\Omega[f(s)] = \alpha(s) \,.\, f(s).$$

It is easily seen that, if $\alpha(s)$ is well-defined, bounded, and summable in (a, b), Ω is precisely distributive in Hilbertian space. Further, if $\alpha(s)$ is continuous in an interval (a', b'), Ω is perfect in this interval. As a second example we may take

$$\Omega[f(s)] = \int_a^b \alpha(s, t) f(t)\, dt.$$

This transformation is precisely distributive in Hilbertian space, if $\{\alpha(s, t)\}^2$ is summable in $a \leq s \leq b$, $a \leq t \leq b$ and $\int_a^b \{\alpha(s, t)\}^2 dt$ exists as a finite number for each value of s in (a, b).[*] If we further suppose that $\alpha(s, t)$ is a continuous function in the rectangle $a' \leq s \leq b'$, $a \leq t \leq b$, the transformation is perfect in the closed interval (a', b'). A wider condition than that just stated may, of course, be given.

The Expansion Theorem.

§ 7. The definition of a function of positive type which was given in the previous communication had reference only to bounded functions.[†] This definition must here be extended.

Let $\gamma(s, t)$ be a symmetric function such that $\{\gamma(s, t)\}^2$ is summable in the square $a \leq s \leq b$, $a \leq t \leq b$. Then, if $f(s)$ is any function which belongs to F, the integral

$$\int_a^b \int_a^b \gamma(s, t) f(s) f(t)\, ds\, dt$$

always exists. We shall say that $\gamma(s, t)$ is of *positive type* in the square if the value of the integral is positive or zero for each function $f(s)$ of the class F.

---

[*] It may be remarked that this transformation has the property that, corresponding to *any* (well defined or not) linear function $\sum_{n=1} x_n f_n(s)$ associated with the class F, the linear function $\sum_{n=1} x_n \Omega[f_n(s)]$ is equal to $\Omega[f((x), s)]$.

[†] *Loc. cit.*, p. 406.

When the integral is positive for each function $f(s)$ which is not a null function, we say that $\gamma(s,t)$ is *definite.*

Let $\theta_1(s), \theta_2(s), \ldots, \theta_n(s), \ldots$ be a countable set of functions which belong to F; and let the series

$$\int_a^b \{\theta_1(s)\}^2\,ds + \int_a^b \{\theta_2(s)\}^2\,ds + \ldots + \int_a^b \{\theta_n(s)\}^2\,ds + \ldots, \tag{9}$$

be convergent. In view of what was said in § 4 and of the inequality

$$\left\{\sum_{n=1}^{r} |\theta_n(s)\,\theta_n(t)|\right\}^2 \leq \sum_{n=1}^{r} \left\{\theta_n(s)\right\}^2 . \sum_{n=1}^{r} \left\{\theta_n(t)\right\}^2,$$

it will be clear that the series $\sum\limits_{n=1} \theta_n(s)\,\theta_n(t)$ is absolutely convergent in the square $a \leq s \leq b$, $a \leq t \leq b$, save possibly at such points as belong to a plane set which has for its projection on either axis a set of measure zero. We may express this more briefly in the statement that the series is absolutely convergent *almost everywhere with respect to s, and almost everywhere with respect to t.* This abbreviated form of statement will be employed under similar circumstances throughout the pages which follow.

From the same inequality and the fact that (9) is convergent, it follows easily that

$$\gamma(s,t) = \theta_1(s)\,\theta_1(t) + \theta_2(s)\,\theta_2(t) + \ldots + \theta_n(s)\,\theta_n(t) + \ldots,$$

has its square summable in $a \leq s \leq b$, $a \leq t \leq b$. Further, if $f(s)$ is any function of F, it would be found that

$$\int_a^b \int_a^b \gamma(s,t) f(s) f(t)\,ds\,dt = \sum_{n=1} \left\{\int_a^b \theta_n(s) f(s)\,ds\right\}^2,$$

and hence that $\gamma(s,t)$ is of positive type in the square.

We observe that, from the hypotheses stated

$$f((x),s) = \sum_{n=1} x_n \theta_n(s)$$

is a linear function associated with the class F, (x) being any real point of Hilbertian space.

§ 8. Now let Ω, Ω' be two linear transformations which form a symmetric couple. It follows from the property of $f((x),s)$ stated at the end of the preceding paragraph that the integral

$$\int_a^b \Omega_s[f((x),s)] \,.\, \Omega_s'[f((x),(s)]\,ds, \tag{10}$$

exists at each point (x). Further, if $f_r((x),s)$ is employed to denote $\sum\limits_{n=1}^{r} x_n\theta_n(s)$, we have

$$\int_a^b \Omega[f_r((x),s)] \,.\, \Omega'[f_r((x),s)]\,ds = \sum_{m=1}^{r}\sum_{n=1}^{r} a_{mn}x_m x_n, \tag{11}$$

where $a_{mn} = \int_a^b \Omega[\theta_m(s)] \, . \, \Omega'[\theta_n(s)] \, ds$ and $a_{mn} = a_{nm}$, because Ω and Ω' form a symmetric couple. Thus, recalling that the left-hand member of (11) tends to (10) as r tends to infinity,[*] we see that $\sum_{m,n} a_{mn}x_m x_n$ is a limited quadratic form,[†] which at each point (x) is represented by (10).

The form $\sum_{m,n} a_{mn}x_m x_n$ may be shown to be both limited and completely continuous. For we have

$$a_{mn}^2 = \left\{ \int_a^b \Omega[\theta_m(s)] \, . \, \Omega'[\theta_n(s)] \, ds \right\}^2$$

$$\leq \int_a^b \{\Omega[\theta_m(s)]\}^2 \, ds \, . \int_a^b \{\Omega'[\theta_n(s)]\}^2 \, ds,$$

whence, using the bounded property of a linear transformation, we have

$$a_{mn}^2 \leq \mathrm{M}^2\mathrm{M}'^2 \int_a^b \{\theta_m(s)\}^2 \, ds \, . \int_a^b \{\theta_n(s)\}^2 \, ds,$$

where the numbers M, M′ are independent of m and n. Since the series

$$\sum_{n=1} \int_a^b \{\theta_n(s)\}^2 \, ds,$$

is convergent by hypothesis, the double series $\sum_{m,n} a_{mn}^2$ is convergent. The convergence of the latter series is known to ensure that the quadratic form $\sum_{m,n} a_{mn}x_m x_n$ is limited and completely continuous.

§ 9. The form $\sum_{m,n} a_{mn}x_m x_n$ being limited and completely continuous, it follows that there exists an orthogonal substitution,

$$x_m' = l_{m1}x_1 + l_{m2}x_2 \ldots + l_{mn}x_n + \ldots (m = 1, 2, \ldots),$$

such that

$$\sum_{m,n} a_{mn}x_m x_n = \sum_{n=1} k_n x_n'^2, \tag{12}$$

where $k_1, k_2, \ldots, k_n$ are constants, some of which may be zero. It will be recalled that the coefficients l of the substitution satisfy the relations

$$\sum_{r=1} l_{mr}l_{nr} = \sum_{r=1} l_{rm}l_{rn} = \begin{cases} 1 \text{ for } m = n, \\ 0 \text{ for } m \neq n, \end{cases} \tag{13}$$

and that the co-ordinates of the point (x) are expressible in terms of those of the point (x') by the equations

$$x_m = l_{1m}x_1' + l_{2m}x_2' + \ldots + l_{nm}x_m' + \ldots (m = 1, 2, \ldots).$$

---

[*] *Vide* § 5.

[†] *Cf.* Hellinger and Toeplitz, 'Math. Annalen,' vol. 69, pp. 326–7.

Employing the equations just written, and recalling that $\sum_{n=1} \{\theta_n(s)\}^2$ converges almost everywhere in (a, b), it will be found that, for almost all values of s in (a, b),

$$f((x), s) = \sum_{n=1} x_n' \chi_n(s), \tag{14}$$

where $X_n(s) = \sum_{m=1} l_{nm}\theta_m(s)$. It is easily seen that $\chi_n(s)$ is a function which belongs to F, and that the series

$$\int_a^b \{\chi_1(s)\}^2 ds + \int_a^b \{\chi_2(s)\}^2 ds + \ldots + \int_a^b \{\chi_n(s)\}^2 ds + \ldots,$$

is convergent.* We conclude that

$$g((x')\, s) = x_1'\chi_1(s) + x_2'\chi_2(s) + \ldots + x_n'\chi_n(s) + \ldots,$$

is a linear function associated with the class F, (x') being a variable point of Hilbertian space.

Recalling that the left-hand member of (12) is (10), we see from (14) that, at each point (x'),

$$\int_a^b \Omega[g((x'), s)] \,.\, \Omega'[g((x'), s)]\, ds = \sum_{n=1} k_n x_n'^2. \tag{15}$$

The left-hand member of this equation can be shown, by an argument similar to that employed in § 8, to represent the quadratic form $\sum_{m, n} b_{mn} x_m' x_n'$, where $b_{mn} = \int_a^b \Omega[\chi_m(s)] \,.\, \Omega'[\chi_n(s)]\, ds$, and we have $b_{mn} = b_{nm}$ in virtue of the hypothesis that Ω and Ω' form a symmetric couple. Comparing the coefficients of like terms on the two sides of (15) we find that

$$\int_a^b \Omega[\chi_m(s)] \,.\, \Omega'[\chi_n(s)]\, ds = \begin{cases} k_m \text{ for } n = m, \\ 0 \text{ for } n \neq m. \end{cases} \tag{16}$$

§ 10. We have seen that the convergence of (9) ensures that $\sum_{n=1} \{\theta_n(s)\}^2$ is convergent almost everywhere. Let us now suppose that (s, t) is a point in the square $a \leq s \leq b$, $a \leq t \leq b$ such that the series $\sum_{n=1} \{\theta_n(s)\}^2$, $\sum_{n=1} \{\theta_n(t)\}^2$ both converge.† Then the series $\sum_{n=1} \theta_n(s)\,\theta_n(t)$ is absolutely convergent and, making

* The sum is equal to that of the series $\sum_{n=1} \int_a^b \{\theta_n(s)\}^2 ds$.

† One, or more, of the functions $\theta_n(s)$, $\theta_n(t)$, $(n = 1, 2, \ldots)$ may be multiple valued at the point under consideration. When this is the case, it must be understood that a function defined in terms of the θ's (*e.g.*, $\gamma(s, t)$ or one of the functions $u_n(s)$, $u_n(t)$ below) is supposed to have the value corresponding to a particular set of determinations $\theta_n(s)$, $\theta_n(t)$, $(n = 1, 2, \ldots)$ at the point. In the pages which follow, when any statement involving several such functions is made, it is understood that each function has the value corresponding to the *same* set of determinations.

use of (13), it will be found that

$$\gamma(s, t) = \sum_{n=1} \theta_n(s) \,.\, \theta_n(t) = \sum_{n=1} \chi_n(s) \,.\, \chi_n(t), \tag{17}$$

where $\chi_n(s)$, $\chi_n(t)$ have finite values for all values of n.

The functions $\chi(s)$ may be divided into two classes according as the corresponding number k is not, or is, zero. Let $u_1(s), u_2(s), \ldots, u_n(s), \ldots$ be the functions $\chi(s)/|k|^{\frac{1}{2}}$ for which the corresponding k is not zero; and let $\mathrm{U}_1(s), \mathrm{U}_2(s), \ldots, \mathrm{U}_n(s) \ldots$ be the functions $\chi(s)$ for which the corresponding k is zero. Further, let μ_n be the reciprocal of the modulus of that k which corresponds to $u_n(s)$; and let δ_n be $+1$ or -1 according as the k referred to is positive or negative. Then from (16) and (17) it will be seen that we have the following theorem:—

Let $\theta_1(s), \theta_2(s), \ldots, \theta_n(s), \ldots$ be a countable set of functions, each of which has its square summable in (a, b); let the series

$$\int_a^b \{\theta_1(s)\}^2\, ds + \int_a^b \{\theta_2(s)\}^2\, ds + \ldots + \int_a^b \{\theta_n(s)\}^2\, ds + \ldots, \tag{8}$$

be convergent; and let $\gamma(s, t)$ be the function of positive type which is defined in the square $a \leqq s \leqq b$, $a \leqq t \leqq b$ by

$$\gamma(s, t) = \sum_{n=1} \theta_n(s) \,.\, \theta_n(t).$$

Further, let Ω, Ω' be a symmetric couple of linear transformations for the interval (a, b). Then, at any point (s, t) of the square such that $\gamma(s, s)$ and $\gamma(t, t)$ are finite, we have*

$$\gamma(s, t) = \sum_{n=1} \frac{u_n(s)\, u_n(t)}{\mu_n} + \sum_{n=1} \mathrm{U}_n(s) \,.\, \mathrm{U}_n(t), \tag{18}$$

where the functions $u_n(s)$, $\mathrm{U}_n(s)$ $(n = 1, 2, \ldots)$ have summable squares in (a, b), the numbers μ_n $(n = 1, 2, \ldots)$ are positive, and each of the series on the right is absolutely convergent. Further, the functions $u(s)$, $\mathrm{U}(s)$ are such that (i)

$$\int_a^b \Omega\,[u_m(s)] \,.\, \Omega'\,[u_n(s)]\, ds = \begin{cases} \delta_m \text{ for } n = m, \\ 0 \text{ for } n \neq m, \end{cases}$$

where δ_m is either $+1$ or -1, and (ii) *each of the pairs of functions* $\{\Omega\,[u_m(s)], \Omega'\,[\mathrm{U}_n(s)]\}$, $\{\Omega\,[\mathrm{U}_m(s)], \Omega'\,[u_n(s)]\}$, $\{\Omega\,[\mathrm{U}_m(s)], \Omega'\,[\mathrm{U}_n(s)]\}$ *is orthogonal for the interval (a, b), whatever be the values of m and n.*

In the theorem, we may suppose $\Omega' = \Omega$. Then $\Omega\,[\mathrm{U}_n(s)]$ is orthogonal to itself, and is therefore a null function for the interval (a, b); further, it will be seen that δ_m is always $+1$, since (10) is, in this case, a positive quadratic form. We have thus the following corollary:—

---

* *Vide* footnote, § 10.

Let Ω *be any linear transformation, and let* $\gamma(s, t)$ *satisfy the hypotheses of the above theorem. Then, under the same conditions as to* $\gamma(s, s)$ *and* $\gamma(t, t)$, *we have* (18) *where* $\Omega[u_n(s)]\,(n = 1, 2, \ldots)$ *is an orthogonal system of functions for the interval* (a, b), *and the functions* $\Omega[\mathrm{U}_n(s)]\,(n = 1, 2, \ldots)$ *are null functions.*

By taking $\Omega = \mathrm{E}$, where E is the unit transformation, and substituting $u_n(s) = \psi_n(s)$, $\mathrm{U}_n(s) = \omega_n(s)$, $\mu_n = \lambda_n$, we obtain the expansion

$$\gamma(s, t) = \sum_{n=1} \frac{\psi_n(s)\,\psi_n(t)}{\lambda_n} + \sum_{n=1} \omega_n(s)\,\omega_n(t). \tag{19}$$

The functions $\psi_n(s)\,(n = 1, 2, \ldots)$ form an orthogonal system of functions for the interval (a, b), and $\omega_n(s)\,(n = 1, 2, \ldots)$ are null functions. It should be recalled that $\gamma(s, t)$ is not necessarily bounded, and that (19) is valid when $\gamma(s, s)$ and $\gamma(t, t)$ are finite. There will be no difficulty in proving that

$$\psi_n(s) = \lambda_n \int_a^b \gamma(s, t)\,\psi_n(t)\,dt,$$

for any value of s such that $\gamma(s, s)$ is finite. Finally, in view of what is said in §§ 14, 15 below, it will be seen that the Fredholm theory can be extended to the function $\gamma(s, t)$.

§ 11. We may employ the corollary of the preceding paragraph to further analyse the expansion of $\gamma(s, t)$. For the function $\sum_{n=1} \mathrm{U}_n(s) \,.\, \mathrm{U}_n(t)$ satisfies the requirements of the corollary.* It follows that we have

$$\sum_{n=1} \mathrm{U}_n(s)\,\mathrm{U}_n(t) = \sum_{n=1} \frac{v_n(s)\,v_n(t)}{\nu_n} + \sum_{n=1} \mathrm{V}_n(s)\,\mathrm{V}_n(t), \tag{20}$$

where $\Omega'[v_n(s)]\,(n = 1, 2, \ldots)$ form an orthogonal system of functions for (a, b), $\nu_n\,(n = 1, 2, \ldots)$ are positive numbers, and $\Omega'[\mathrm{V}_n(s)]\,(n = 1, 2, \ldots)$ are null functions. As the functions $v_n(s)$, $\mathrm{V}_n(s)$ are linear functions of $\mathrm{U}_n(s)\,(n = 1, 2, \ldots)$, it will be clear that the orthogonal properties stated under (ii) in the theorem of the preceding paragraph will hold when the function symbol, U, is replaced by either v or V. There is, however, no need to state the orthogonal properties which involve functions $\Omega'[\mathrm{V}_n(s)]$, since these are a consequence of the fact that $\Omega'[\mathrm{V}_n(s)]\,(n = 1, 2, \ldots)$ are null functions.

Next, we can apply the same corollary to $\Sigma\ \mathrm{V}_n(s)\,\mathrm{V}_n(t)$ and so obtain

$$\sum_{n=1} \mathrm{V}_n(s)\,\mathrm{V}_n(t) = \sum_{n=1} \frac{w_n(s)\,.\,w_n(t)}{\rho_n} + \sum_{n=1} \mathrm{W}_n(s)\,\mathrm{W}_n(t), \tag{21}$$

where $\Omega[w_n(s)]\,(n = 1, 2, \ldots)$ form an orthogonal system of functions for (a, b), $\rho_n\,(n = 1, 2, \ldots)$ are positive numbers, and $\Omega[\mathrm{W}_n(s)]\,(n = 1, 2, \ldots)$ are null functions. After what was said above, the reader will have no difficulty in perceiving that the orthogonal properties are as stated in the theorem which follows.

---

* The hypothesis that $\gamma(s, s)$ and $\gamma(t, t)$ are finite is still understood.

Employing (20) and (21), it will be seen that we have the following theorem :—

Under the hypotheses of the theorem of the last paragraph, we have

$$\gamma(s, t) = \sum_{n=1} \frac{u_n(s)\, u_n(t)}{\mu_n} + \sum_{n=1} \frac{v_n(s)\, v_n(t)}{\nu_n}$$
$$+ \sum_{n=1} \frac{w_n(s)\, w_n(t)}{\rho_n} + \sum_{n=1} \mathrm{W}_n(s) \,.\, \mathrm{W}_n(t), \quad (22)$$

where the functions $u(s)$, $v(s)$, $w(s)$, $\mathrm{W}(s)$ *have summable squares in* (a, b), *the numbers* μ, ν, ρ *are positive, and each of the series on the right is absolutely convergent. Further, the functions* $u(s)$, $v(s)$, $w(s)$, $\mathrm{W}(s)$ *are such that* (i)

$$\int_a^b \Omega[u_m(s)] \,.\, \Omega'[u_n(s)]\, ds = \begin{cases} \delta_m \text{ for } n = m, \\ 0 \text{ for } n \neq m, \end{cases}$$

where δ_m *is either* $+1$ *or* -1; (ii) *the functions* $\Omega'[v_n(s)]$ $(n = 1, 2, \ldots)$ *and the functions* $\Omega[w_n(s)]$ $(n = 1, 2, \ldots)$ *each constitute an orthogonal system of functions for the interval* (a, b); (iii) *each of the pairs of functions* $\{\Omega[u_m(s)], \Omega'[v_n(s)]\}$, $\{\Omega[v_m(s)], \Omega'[u_n(s)]\}$, $\{\Omega[v_m(s)], \Omega'[v_n(s)]\}$, $\{\Omega[w_m(s)], \Omega'[u_n(s)]\}$, $\{\Omega[w_m(s)], \Omega'[v_n(s)]\}$, *is orthogonal for* (a, b), *whatever be the values of* m *and* n; *and* (iv) $\Omega'[w_n(s)]$, $\Omega[\mathrm{W}_n(s)]$, $\Omega'[\mathrm{W}_n(s)]$ *are null functions.*

The function $\gamma(s, s)$ has a single finite value almost everywhere in (a, b). It follows that (22) is valid almost everywhere with respect to s, and almost everywhere with respect to t.

It is perhaps unnecessary to point out that there is a second expansion of $\gamma(s, t)$ of the form (22); for this expansion, the statements of the above theorem are applicable, provided that Ω and Ω' are everywhere interchanged.* In either case, one or more of the series on the right of (22) may contain only a finite number of terms, and it is possible that one or more of the series may be non-existent.

§ 12. Referring to the theorem of § 11, let us denote the sum function of the series $\sum_{n=1} u_n(s)\, u_n(t)/\mu_n$ by $\gamma_1(s, t)$. Since $\gamma(s, s)$ is summable in (a, b), it will be clear that $\sum_{n=1} \int_a^b \{u_n(s)/\mu_n^{\frac{1}{2}}\}^2 ds$ is convergent; and hence that

---

* To derive this expansion Ω and Ω' must, of course, be interchanged in the argument at the commencement of the paragraph.

$\sum\limits_{n=1} x_n u_n(s)/\mu_n^{\frac{1}{2}}$ is a linear function associated with F. It follows from what was said in § 5, that

$$\Omega_s\Big[\sum_{n=1} x_n u_n(s)/\mu_n^{\frac{1}{2}}\Big] = \sum_{n=1} x_n \Omega[u_n(s)]/\mu_n^{\frac{1}{2}}$$

almost everywhere in (a, b). Observing that $(u(t)/\mu^{\frac{1}{2}})$ is a point of Hilbertian space for almost all values of t in the interval (a, b), we see that

$$\Omega_s[\gamma_1(s, t)] = \sum_{n=1} \Omega[u_n(s)]\, u_n(t)/\mu_n, \tag{23}$$

almost everywhere with respect to s, and almost everywhere with respect to t.

Again we have

$$\Omega_t'\Big[\sum_{n=1} x_n u_n(t)/\mu_n^{\frac{1}{2}}\Big] = \sum_{n=1} x_n \Omega_t'[u_n(t)]/\mu_n^{\frac{1}{2}}$$

almost everywhere in (a, b) and $(\Omega[u(s)]/\mu^{\frac{1}{2}})$ is a point of Hilbertian space for almost all values of s in the interval $a \leq s \leq b$. It follows that we have

$$\Omega_t'\Big[\sum_{n=1} \Omega[u_n(s)] \,.\, u_n(t)/\mu_n\Big] = \sum_{n=1} \Omega[u_n(s)] \,.\, \Omega'[u_n(t)]/\mu_n$$

almost everywhere with respect to s, and almost everywhere with respect to t. We conclude from (23) that

$$\Omega_t'[\Omega_s[\gamma_1(s, t)]] = \sum_{n=1} \Omega[u_n(s)]\, \Omega'[u_n(t)]/\mu_n, \tag{24}$$

almost everywhere with respect to s, and almost everywhere with respect to t.

It will be clear that in a similar way we can prove the validity of

$$\Omega_s[\Omega_t'[\gamma_1(s, t)]] = \sum_{n=1} \Omega[u_n(s)] \,.\, \Omega'[u_n(t)]/\mu_n, \tag{25}$$

under similar circumstances. Adopting a slightly abbreviated mode of expression for the left-hand members of (24) and (25), we see that

$$\Omega_s\Omega_t'[\gamma_1(s, t)] = \Omega_t'\Omega_s[\gamma_1(s, t)] = \sum_{n=1} \Omega[u_n(s)] \,.\, \Omega'[u_n(t)]/\mu_n$$

almost everywhere with respect to s, and almost everywhere with respect to t.

Similar results may be established in respect to the other series on the right of (22), and hence in respect to $\gamma(s, t)$. Further, it will be clear that each series such as that on the right of (25) is absolutely convergent almost everywhere with respect to s, and almost everywhere with respect to t.

§ 13. Let us define a function $\kappa(s, t)$ by the rule that it has the common value of $\Omega_s\Omega'_t[\gamma(s, t)]$ and $\Omega'_t\Omega_s[\gamma(s, t)]$ at points where these are equal. Further, let

$$\begin{aligned} \phi_n(s) &= \delta_n \,.\, \Omega[u_n(s)], & \psi_n(s) &= \Omega'[u_n(s)], \\ \xi_n(s) &= \Omega[v_n(s)], & \eta_n(s) &= \Omega'[v_n(s)], \\ \zeta_n(s) &= \Omega[w_n(s)], & \omega_n^{(1)}(s) &= \Omega'[w_n(s)], \\ \omega^{(2)}(s) &= \Omega[W_n(s)], & \omega_n^{(3)}(s) &= \Omega'[W_n(s)], \end{aligned}$$

and let $\lambda_n = \delta_n \mu_n$. It will be seen that we have

$$\kappa(s, t) = \sum_{n=1} \frac{\phi_n(s)\psi_n(t)}{\lambda_n} + \sum_{n=1} \frac{\xi_n(s)\eta_n(t)}{\upsilon_n},$$

$$+ \sum_{n=1} \frac{\zeta_n(s)\omega_n^{(1)}(t)}{\rho_n} + \sum_{n=1} \omega_n^{(2)}(s)\,\omega_n^{(3)}(t),^* \qquad (26)$$

almost everywhere with respect to s, and almost everywhere with respect to t; also it can be established that term by term integration with respect to either variable of the series on the right is justified. From these facts and the properties of the functions $\phi_n(s)$, $\psi_n(s)$, ..., which follow from the statements (i)–(iv) in the theorem of § 11, we find that, almost everywhere in (a, b), we have

$$\left.\begin{aligned} \phi_n(s) &= \lambda_n \int_a^b \kappa(s, t)\,\phi_n(t)\,dt \\ \psi_n(t) &= \lambda_n \int_a^b \psi_n(s)\,\kappa(s, t)\,ds \end{aligned}\right\}. \qquad (27)$$

We can now state the following theorem:—

Let $\gamma(s, t)$ be a function of positive type defined as in the theorem of § 10; *let Ω, Ω' be a symmetric couple of linear transformations; and let $\kappa(s, t)$ be the common value of $\Omega_s\Omega'_t[\gamma(s, t)]$ and $\Omega'_t\Omega_s[\gamma(s, t)]$ at points where these are equal. Then the equation* (26) *is valid, and the right-hand member is absolutely convergent, almost everywhere with respect to s, and also everywhere with respect to t. The functions*

$$\phi_1(s),\ \phi_2(s),\ \ldots,\ \phi_n(s),\ \ldots,$$
$$\psi_1(s),\ \psi_2(s),\ \ldots,\ \psi_n(s),\ \ldots,$$

form a bi-orthogonal system of functions for the interval (a, b), and each of the homogeneous equations (27) *is satisfied almost everywhere in the interval. The symbols υ_n, ρ_n $(n = 1, 2, \ldots)$ denote positive constants. The functions $\eta_n(s)$ $(n = 1, 2, \ldots)$ and the functions $\zeta_n(s)$ $(n = 1, 2, \ldots)$ each constitute an orthogonal system of functions for the interval (a, b). Each of the pairs of functions $\{\phi_m(s), \eta_n(s)\}$, $\{\xi_m(s), \psi_n(s)\}$, $\{\xi_m(s), \eta_n(s)\}$, $\{\zeta_m(s), \psi_n(s)\}$ $\{\zeta_m(s), \eta_n(s)\}$, is orthogonal for the interval (a, b), whatever be the values of m and n. Lastly, for all values of n the functions $\omega_n^{(1)}(s)$, $\omega_n^{(2)}(s)$, $\omega_n^{(3)}(s)$ are null functions.*

In view of a remark at the end of § 11, it will be seen that there is a second expansion for $\kappa(s, t)$, similar to the above.

---

* For the case under consideration, the series $\sum_{n=1} \omega^{(2)}(s), \omega^{(3)}(t)$ may be omitted, since its sum function is zero almost everywhere with respect to s, and almost everywhere with respect to t. The series is retained here in view of the expansion of $\gamma(s, t)$, and because it cannot be ignored when the transformations are precisely distributive in Hilbertian space (*vide* § 17).

§ 14. It seems desirable to make some observations in regard to the theorem of the last paragraph.

In the first place, it will be clear that the Fredholm theory, as usually developed, is not applicable to $\kappa(s, t)$, since this function may be unbounded in such a general manner. We may, however, obtain results similar to those of the theory referred to, as follows:—

If we define $K(\lambda; s, t)$ to be the sum function of the series on the right of (26) after λ_n has been replaced by $\lambda_n - \lambda$ in the first component series, we find that the square of $K(\lambda; s, t)$ is summable in $a \leq s \leq b$, $a \leq t \leq b$, provided that λ has not one of the values $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$.

It is easily verified that

$$K(\lambda; s, t) - \kappa(s, t) = \lambda \int_a^b K(\lambda; s, u)\, \kappa(u, t)\, du = \lambda \int_a^b \kappa(s, u)\, K(\lambda; u, t)\, du, \tag{28}$$

almost everywhere with respect to s, and almost everywhere with respect to t; and hence that, if $f(s)$ belongs to the class F, the equation

$$f(s) = \phi(s) - \lambda \int_a^b \kappa(s, t)\, g(t)\, dt, \tag{29}$$

is satisfied almost everywhere in (a, b) by

$$g(s) = f(s) + \lambda \int_a^b K(\lambda; s, t)\, f(t)\, dt,$$

provided that λ is not one of the numbers λ_n. It will be clear that a corresponding remark applies in regard to the solution of the equation adjoint to (29).

The numbers $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ may be called *singular values* of $\kappa(s, t)$.*

§ 15. Referring to §§ 8, 9, it will be found that

$$|k_n| = \left| \int_a^b \Omega[\chi_n(s)] \,.\, \Omega'[\chi_n(s)]\, ds \right| \leq MM' \int_a^b \{\chi_n(s)\}^2 ds.$$

Since $\sum\limits_{n=1} \{\chi_n(s)\}^2 = \sum\limits_{n=1} \{\theta_n(s)\}^2$ almost everywhere, we see that

$$\sum_{n=1}^{m} |k_n| \leq MM' \int_a^b \sum_{n=1} \{\theta_n(s)\}^2 ds.$$

---

* The reader will observe that by a known property of completely continuous quadratic forms, $\lim\limits_{n \to \infty} k_n = 0$, and hence that $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$, can be divided into sets, each containing a finite number of equal singular values, and such that members of distinct sets are unequal.

It follows that the series

$$\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\ldots+\frac{1}{\lambda_n}+\ldots, \tag{30}$$

is absolutely convergent, and that $\Sigma\,|1/\lambda_n|$ is $\leq \mathrm{MM}'\int_a^b \gamma(s,s)\,ds$.*

The series (30) being absolutely convergent, it follows that

$$\mathrm{D}(\lambda) = \prod_{n=1}\left(1-\frac{\lambda}{\lambda_n}\right)$$

is an integral function of λ. Also from the expansion of $\kappa(s,t)$, it will be found that, for all integral values of r,

$$\sum_{n=1} \frac{1}{\lambda_n{}^r} = \int_a^b \kappa_r(s,s)\,ds,$$

where $\kappa_1(s,t) = \kappa(s,t)$ and $\kappa_{r+1}(s,t) = \int_a^b \kappa_r(s,u)\,\kappa(u,t)\,du$.

From these facts it can be shown that

$$\mathrm{D}(\lambda) = 1-\lambda\int_a^b \kappa(s_1,s_1)\,ds_1 + \frac{\lambda^2}{2!}\int_a^b\int_a^b \kappa\begin{pmatrix} s_1\, s_2 \\ s_1\, s_2\end{pmatrix} ds_1\,ds_2$$
$$-\ldots+(-)^n\frac{\lambda^n}{n!}\int_a^b\ldots\int_a^b \kappa\begin{pmatrix} s_1, s_2, \ldots, s_n \\ s_1, s_2, \ldots, s_n\end{pmatrix} ds_1\,ds_2\ldots ds_n+\ldots,$$

as in the usual Fredholm theory. Observing that $\mathrm{D}(\lambda)\,\mathrm{K}(\lambda;\, s, t)$ is an integral function of λ, and making use of (28) it can be shown that, almost everywhere with respect to s and almost everywhere with respect to t, we have

$$\mathrm{D}(\lambda)\,\mathrm{K}(\lambda;\, s, t) = \mathrm{D}(\lambda;\, s, t),$$

where $\mathrm{D}(\lambda;\, s, t)$ is the integral function of λ defined by

$$\kappa(s,t)-\lambda\int_a^b \kappa\begin{pmatrix} s\, s_1 \\ t\, s_1\end{pmatrix} ds_1+\ldots+(-)^n\frac{\lambda^n}{n!}\int_a^b\ldots\int_a^b \kappa\begin{pmatrix} s, s_1, s_2, \ldots s_n \\ t, s_1, s_2, \ldots s_n\end{pmatrix} ds_1\,ds_2\ldots ds_n+\ldots.$$

The function $\mathrm{K}(\lambda;\, s, t)$ is thus expressed as the quotient of two integral functions of λ, for almost all values of s, and almost all values of t.

Further developments in this direction may be left to the reader.

§ 16. It can be proved that, if the equation

$$f(s) = \lambda\int_a^b \kappa(s,t)\,f(t)\,dt$$

---

* This is a generalisation of a theorem given by Garbe ('Math. Annalen,' vol. 76, pp. 545–6) for Hilbert's function $a(s)\,\gamma(st)$, where $\alpha(s) = +1$ in certain sub-intervals of (a, b), and $= -1$ in the complementary set. The absolute convergence may also be inferred from the fact (*vide* below) that $\Sigma 1/\lambda_n = \int_a^b \kappa(s,s)\,ds$, and that the order of the singular values is arbitrary.

is satisfied almost everywhere in (a, b) by a function $f(s)$, which is not a null-function, then λ must be one of the singular values λ_n; also $f(s)$ must, almost everywhere in (a, b), be a linear combination of the functions $\phi_n(s)$ corresponding to the singular value.

Again, if $\Gamma(s, t) = \Omega_s' \, \Omega_t' \, [\gamma(s, t)]$, it will be found that

$$\int_a^b \Gamma(s, u)\,\kappa(u, t)\,du = \sum_{n=1} \frac{\psi_n(s)\,\psi_n(t)}{\lambda_n \mu_n},$$

almost everywhere with respect to s, and almost everywhere with respect to t; moreover $\int_a^b \Gamma(s, u)\,\phi_n(u)\,du$ being almost everywhere equal to $\psi_n(s)/\mu_n$ is never a null-function. Giving an extended meaning to the term as previously used,* we may, therefore, say that $\kappa(s, t)$ is completely symmetrisable on the left by $\Gamma(s, t)$. In a similar sense, $\kappa(s, t)$ is completely symmetrisable on the right by $\Gamma'(s, t) = \Omega_s \Omega_t [\gamma(s, t)]$.

Finally, if the equation (26) is multiplied through by $f(s)\,g(t)$, two functions which belong to F, we find after integration that

$$\int_a^b \int_a^b \kappa(s, t) f(s)\,g(t)\,ds\,dt = \sum_{n=1} \frac{\int_a^b \phi_n(s)\,f(s)\,ds \,.\, \int_a^b \psi_n(t)\,g(t)\,dt}{\lambda_n} + \sum_{n=1} \frac{\int_a^b \xi_n(s) f(s)\,ds \,.\, \int_a^b \eta_n(t)\,g(t)\,dt}{\nu_n}. \qquad (31)$$

This result is, of course, analogous to Hilbert's expansion of

$$\int_a^b \int_a^b \kappa(s, t) f(s)\,g(t)\,ds\,dt,$$

for the case when $\kappa(s, t)$ is symmetric.

§ 17. Let $\gamma(s, t)$ be a function which is bounded, summable, and of positive type in the square $a \leq s \leq b$, $a \leq t \leq b$; further, let the function be continuous with respect to s in the interval (a, b) for each fixed value of t in this interval. The fact that $\gamma(s, t)$ is symmetric permits the interchange of s and t in the last statement. Then it is known that†

$$\gamma(s, t) = \sum_{n=1} \frac{\pi_n(s) \,.\, \pi_n(t)}{\sigma_n},$$

where $\pi_n(s)$ $(n = 1, 2, \ldots)$ constitute a complete system of fundamental functions of $\gamma(s, t)$, and σ_n is the (positive) singular value corresponding to

---

* *Cf.* 'Roy. Soc. Proc.,' A, vol. 97, p. 402. The continuity properties assumed in this paper excluded the case of a null-function which is not zero everywhere.

† *Vide* 'Roy. Soc. Proc.,' A, vol. 97, p. 407.

$\pi_n(s)$. The functions $\pi_n(s)$ are continuous in (a, b) and the series $\sum_{n=1} \{\pi_n(s)\}^2/\sigma_n$ is convergent for each value of s which is $\geq a$ and $\leq b$. It follows that, if we write $\pi_n(s)/\sigma_n^{\frac{1}{2}} = \theta_n(s)$, we have

$$\gamma(s, t) = \sum_{n=1} \theta_n(s)\,\theta_n(t),$$

where the functions $\theta_n(s)$ satisfy the requirements of the theorem of § 10. The expansion (22) is thus seen to be applicable at *any* point of the square $a \leq s \leq b$, $a \leq t \leq b$.

Let us now suppose that the transformations Ω, Ω' are well defined and precisely distributive in Hilbertian space, and that, taken together, they constitute a symmetric couple. Proceeding by an argument parallel to that of § 12, it will be found that we have

$$\Omega_s\Omega_t'[\gamma(s, t)] = \Omega_t'\Omega_s[\gamma(s, t)]$$

at *all* points; further, if the common value is denoted by $\kappa(s, t)$, the expansion (26) is valid at *all* points of the square $a \leq s \leq b$, $a \leq t \leq b$.

There can be no difficulty in developing a theory of the integral equation with $\kappa(s, t)$ as kernel on the lines sketched in §§ 14, 15. The reader will observe the essential difference that statements, which in the paragraphs referred to are true almost everywhere, are now true at *all* points.

The functions $\theta_n(s) = \pi_n(s)/\sigma_n^{\frac{1}{2}}$ being all continuous in (a, b), it follows that the functions $u(s)$, $v(s)$, $w(s)$, $W(s)$ are all continuous in the interval. Hence, if the transformation Ω' is perfect in an interval (a', b'), where $a - a' \leq b' \quad b$, the functions $\psi_n(s)$, $\eta_n(s)$, $\omega_n^{(1)}(s)$, $\omega^{(3)}(s)$ are all continuous in (a', b'). Since a continuous null-function is zero everywhere, we see that $\omega_n^{(1)}(s)$, $\omega_n^{(3)}(s)$ are evanescent.

§ 18. We may now state the following expansion theorem :—

Let $\gamma(s, t)$ be a function which is bounded, summable, and of positive type, in the square $a \leq s \leq b$, $a \leq t \leq b$; further, let the function be continuous with respect to s in the interval (a, b) for each fixed value of t in the same interval. Finally, let Ω and Ω' be two linear transformations which are precisely distributive in Hilbertian space, and which together constitute a symmetric couple. Then, at all points of the square $a \leq s \leq b$, $a \leq t \leq b$, we have $\Omega_s\Omega_t'[\gamma(s, t)] = \Omega_t'\Omega_s[\gamma(s, t)]$, and, denoting the common value of these functions by $\kappa(s, t)$, we have the expansion (26), wherein each of the series on the right is absolutely convergent. The functions

$$\phi_1(s),\ \phi_2(s),\ \ldots,\ \phi_n(s), \ldots\ldots$$

$$\psi_1(s),\ \psi_2(s),\ \ldots,\ \psi_n(s), \ldots\ldots$$

which occur in the expansion form a bi-orthogonal system of functions for the interval (a, b) *; and* $\phi_n(s)$, $\psi_n(s)$ *satisfy the equations*

$$\phi_n(s) = \lambda_n \int_a^b \kappa(s, t)\, \phi_n(t)\, dt,$$

$$\psi_n(t) = \lambda_n \int_a^b \psi_n(s)\, \kappa(s, t)\, ds,$$

at all points of the interval (a, b). *The remaining functions which occur in the expansion have the properties stated in the theorem of* § 13.

Moreover, if Ω' *is perfect in an interval* (a', b') *contained in* (a, b), *the expansion takes the form*

$$\kappa(s, t) = \sum_{n=1} \frac{\phi_n(s)\,\psi_n(t)}{\lambda_n} + \sum_{n=1} \frac{\xi_n(s)\,\eta_n(t)}{\nu_n}, \tag{32}$$

for $a \leq s \leq b$, $a' \leq t \leq b'$ *; and the functions* $\psi_n(s)$, $\eta_n(s)$ *are continuous in* $(a'\ b')$. *The corresponding statement, which applies when* Ω *is perfect in a sub-interval of* (a, b), *will be obvious.*

The reader will recall that, in any given case, there is always a second expansion for $\kappa(s, t)$ (*vide* §§ 11, 13).

§ 19. In conclusion, it should be observed that conditions may be stated under which the expansions given in §§ 13, 18 assume a special character. The conditions are generalisations of those stated at the end of § 9 of the communication of April 12 for the case of the function of $\alpha(s)\,\gamma(s, t)$.

For example, let the hypotheses of the theorem of § 13 be adopted, and let $\Gamma(s, t) = \Omega_s'\Omega_t'[\gamma(s, t)]$ be definite. Employing the equation (31) of § 16, it is easily shown that $\xi_n(s)$ $(n = 1, 2, \ldots)$ are null functions, and that $\zeta_n(s)$ $(n = 1, 2, \ldots)$ cannot exist. It follows that the expansion (26) takes the simpler form

$$\kappa(s, t) = \sum_{n=1} \frac{\phi_n(s)\,\psi_n(t)}{\lambda_n} + \sum_{n=1} \frac{\omega_n^{(1)}(s)\,\eta_n(t)}{\nu_n} + \sum_{n=1} \omega_n^{(2)}(s) \cdot \omega_n^{(3)}(t),$$

where the functions which occur have the properties stated in the theorem of § 13. This expansion will hold at all points of $a \leq s \leq b$, $a \leq t \leq b$, when $\gamma(s, t)$ satisfies the requirements of the theorem of the preceding paragraph and, as there, Ω, Ω' are precisely distributive in Hilbertian space. If we further suppose that Ω is perfect in an interval (a', b'), $\omega_n^{(1)}(s)$, $\omega_n^{(2)}(s)$ are continuous null functions, and therefore vanish everywhere in the interval. We thus have the expansion

$$\kappa(s, t) = \sum_{n=1} \frac{\phi_n(s)\,\psi_n(t)}{\lambda_n},$$

which is valid when s belongs to (a', b') and t to (a, b).

It should be observed that the hypothesis that $\gamma(s, t)$ is definite does not in itself ensure that $\Gamma(s, t)$ is definite. This, however, will certainly be the case when Ω' has an inverse.

The Dissipation of Energy in Permanent Ocean Currents, with some Relations between Salinities, Temperatures and Currents.

By R. O. Street, M.A., M.Sc., Fellow of St. John's College, Cambridge, Lecturer in Mathematics and Mathematical Hydrography in the University of Liverpool.

(Communicated by Sir Joseph Larmor, F.R.S. Received January 27, 1921.)

In a previous paper* the author has discussed the effect of viscosity in dissipating the energy of non-turbulent tidal flow of a sea, and the present one contains a simple extension of the method to non-periodic motion. In practically all parts of the open ocean there exist currents whose directions are almost constant and whose velocities may therefore be supposed to contain a component which is independent of the time. The motion is usually slow, the current seldom exceeding one or two knots, and on the average much less, so that in all probability the effect of turbulence will not be large. The work is therefore carried out on the assumption of a smooth flow, and a first approximation only is made. The results are used to obtain an estimate of the mean rate at which energy is dissipated in all the oceans by these permanent current streams.

The second portion of the paper is devoted to a simple method of deducing the magnitude of the surface current solely from a knowledge of the distribution of the salinity or of the temperature of the water. There are not sufficient data available to obtain exact figures, but rough estimates of the orders of magnitude lead to concordant results.

The Dissipation of Energy in Permanent Ocean Currents.

Consider a portion of the sea so small that in the state of relative equilibrium its surface may be regarded as flat, and let it rotate about the vertical with an angular velocity ω which is to be taken as the component of the earth's diurnal rotation. Let the axis of rotation be taken as the axis of z, the origin being at the bottom of the sea and the equation of the mean surface being $z = h$. Take axes of x, y, rotating in their plane with the prescribed angular velocity ω, and let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ be the components referred to these axes of the small relative velocity of the particle of water which at time t is at the point (x, y, z).

---

* 'Roy. Soc. Proc.,' A, vol. 93, p. 348 (1917).

For the horizontal motion the equations are*

$$\left.\begin{aligned}\frac{\partial \mathbf{U}}{\partial t}-2\omega\mathbf{V} &= -g\frac{\partial \mathbf{Z}}{\partial x}+\nu\Delta\mathbf{U}\\ \frac{\partial \mathbf{V}}{\partial t}+2\omega\mathbf{U} &= -g\frac{\partial \mathbf{Z}}{\partial y}+\nu\Delta\mathbf{V}\end{aligned}\right\}, \tag{1}$$

where the density is assumed uniform, ν is the kinematic coefficient of viscosity, Δ is the Laplacian operator, and $\mathbf{Z}-\Pi/g$ is the elevation of the surface above its position of relative equilibrium, Π being the disturbing potential.

Let
$$\left.\begin{aligned}\mathbf{U} &= \mathrm{U}+\Sigma\, u_\sigma \cos(\sigma t-\alpha)\\ \mathbf{V} &= \mathrm{V}+\Sigma\, v_\sigma \cos(\sigma t-\beta)\end{aligned}\right\}, \tag{2}$$

where U, V, u_σ, v_σ are independent of t, and the summation includes terms of different periods. The parts of the solutions which are periodic functions have been discussed previously† on the assumption σ and ω are of the same order of magnitude and that $\sigma-2\omega$ is positive. By a very slight modification the work can be extended to the cases in which σ is less than 2ω, that is, to the longer period motion. We proceed to consider the parts U, V of the solutions which are independent of the time, assuming that ω is not zero.

By equations (1) and (2) these are given by

$$\left.\begin{aligned}-2\omega\mathrm{V} &= -g\frac{\partial \mathrm{Z}}{\partial x}+\nu\frac{\partial^2 \mathrm{U}}{\partial z^2}\\ 2\omega\mathrm{U} &= -g\frac{\partial \mathrm{Z}}{\partial y}+\nu\frac{\partial^2 \mathrm{V}}{\partial z^2}\end{aligned}\right\}, \tag{3}$$

where Z is the part of $\mathbf{Z}$ independent of the time, and $\Delta(\mathrm{U}, \mathrm{V})$ have been replaced by $\partial^2(\mathrm{U}, \mathrm{V})/\partial z^2$.

Combining these equations we obtain

$$\left(\frac{\partial^2}{\partial z^2}\mp\frac{2\iota\omega}{\nu}\right)(\mathrm{U}\pm\iota\mathrm{V}) = \frac{g}{\nu}\left(\frac{\partial}{\partial x}\pm\iota\frac{\partial}{\partial y}\right)\mathrm{Z},$$

of which the solutions, subject to the conditions U, V zero at the bottom $z=0$, and $\partial\mathrm{U}/\partial z$, $\partial\mathrm{V}/\partial z$ zero at the surface $z=h$, are

$$\mathrm{U}\pm\iota\mathrm{V} = \pm\frac{\iota g}{2\omega}\left[1-\frac{\cosh\{\gamma(h-z)(1\pm\iota)\}}{\cosh\{\gamma h(1\pm\iota)\}}\right]\left(\frac{\partial}{\partial x}\pm\iota\frac{\partial}{\partial y}\right)\mathrm{Z}, \tag{4}$$

where $\gamma^2=\omega/\nu$.

In the open sea γh is always very large, so that we can replace these solutions by

$$\mathrm{U}\pm\iota\mathrm{V} = (\mathrm{U}_0\pm\iota\mathrm{V}_0)\left[1-e^{-\gamma z(1\pm\iota)}\right], \tag{5}$$

where U_0, V_0 are the surface values of U, V.

---

* See, for example, Lamb, 'Hydrodynamics,' §§ 206, 316.

† 'Roy. Soc. Proc.,' A, *loc. cit.*

Now let us suppose that Q is the local magnitude of the non-periodic part of the surface current relative to the earth's rotation, and that the local axes are chosen so that Q is in the direction Ox. Then $U_0 = Q$, $V_0 = 0$, and equations (5) become

$$\left.\begin{aligned} U &= Q\left(1-e^{-\gamma z}\cos\gamma z\right) \\ V &= Qe^{-\gamma z}\sin\gamma z \end{aligned}\right\}. \qquad (6)$$

It is now easy to justify the neglect of the terms in obtaining equations (3). The non-periodic terms in the equation of continuity lead to the relation

$$\int_0^h \left(\frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}\right)dz = 0,$$

which from equations (4) is equivalent to

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)Z = 0.$$

Hence from the same equations

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)(U, V) = 0,$$

so that the neglected terms are, to this approximation, actually zero.

The total rate of dissipation of energy per unit surface area of the sea may be taken* as

$$-\mu\int_0^h \left(\mathbf{U}\frac{\partial^2\mathbf{U}}{\partial z^2}+\mathbf{V}\frac{\partial^2\mathbf{V}}{\partial z^2}\right)dz$$

where μ is the coefficient of viscosity.

Substituting from equations (2) and taking the mean value over a long period of time we obtain a result which can be expressed in the form

$$-\mu\int_0^h \left(U\frac{\partial^2 U}{\partial z^2}+V\frac{\partial^2 V}{\partial z^2}\right)dz+\Sigma E_\sigma,$$

where E_σ is the contribution from the terms of period $2\pi/\sigma$. The mean rate of energy dissipation due to each term in the current velocity can thus be evaluated separately and the several results added. The contribution from the non-periodic terms given by equations (6) is found to be $\mu\gamma Q^2$ to the order of approximation already used, thus depending only on the velocity and not involving its gradients.

Writing $\omega = \Omega\sin\lambda$, where Ω is the earth's diurnal rotation and λ is the latitude of the place, we obtain for the mean rate, F, of dissipation of energy per unit surface area due to the permanent ocean currents

$$F \doteqdot \rho Q^2\sqrt{(\nu\Omega\sin\lambda)}, \qquad (7)$$

---

[*] R. O. Street, *loc. cit.*, p. 352.

where ρ is the density of the water* and ν is the kinematic coefficient of viscosity appropriate to the circumstances.

For places on the equator ω is zero and equations (3) reduce to

$$\frac{\partial^2 U}{\partial z^2} = \frac{g}{\nu}\frac{\partial Z}{\partial x}, \qquad \frac{\partial^2 V}{\partial z^2} = \frac{g}{\nu}\frac{\partial Z}{\partial y},$$

so that, with the same surface and bottom conditions as before, we have in place of (5)

$$\frac{U}{U_0} = \frac{V}{V_0} = \frac{2z}{h} - \frac{z^2}{h^2}.$$

Hence we obtain in place of equation (7)

$$F' = \frac{4}{3}\frac{\mu Q^2}{h}. \tag{8}$$

Now in deep water F′ is very small compared with $\rho Q^2 \sqrt{(\nu\Omega)}$. Therefore in estimating the mean rate at which energy is dissipated by the permanent currents of all the oceans we may without appreciable error suppose that equation (7) is valid everywhere.

In order to obtain a numerical estimate for the total actual rate of energy dissipation due to these non-periodic currents, the surface of the sea has been divided into 10-degree "squares" and for each a mean value of Q has been used. These values have been deduced from the Admiralty current charts. At any point the speed of the current is usually variable; for the non-periodic part, the mean of the two extreme values has been taken. Thus when the current is recorded as 20–40 sea miles per day the value 30 has been used. The variations in current strengths are greatest in the neighbourhood of the equator; owing to the presence of the factor $\sin\lambda$ in the formula for the rate of energy dissipation the effect of this variation is not important. In any 10-degree "square" the mean latitude is taken as sufficiently accurate for our present purpose.

The details of the arithmetic are omitted and only the final results for each of the three great oceans given.

For the Atlantic, Pacific, and Indian Oceans the values

$$(7,\ 11,\ 7) \times 10^{17} \text{ ergs per second}, \tag{9}$$

have been obtained by substituting in the formula (7), and assuming that $\nu = 0{\cdot}018$, the areas over which the summation has been carried out being respectively

$$(7{\cdot}3,\ 13{\cdot}5,\ 5{\cdot}4) \times 10^{17} \text{ sq. cm.}$$

---

* In this portion of the work no appreciable error is introduced by regarding ρ as constant, although in all probability the variations of density actually constitute a primary cause of the currents.

The areas of the surfaces of these oceans are

$$(8{\cdot}2,\ 16{\cdot}6,\ 7{\cdot}3) \times 10^{17} \text{ sq. cm.}$$

but there is little information available beyond the fiftieth parallel of latitude, or the sixtieth in the case of the North Atlantic. Hence we may say that, on the explicit assumption of non-turbulent flow, the presence of viscosity causes the non-periodic parts of the ocean currents to produce a dissipation of energy at the mean rate

$$3 \times 10^{18} \text{ ergs per second.} \qquad (10)$$

It may be noted that this result is roughly 1 erg per second per square centimetre of surface. The estimate (10) is to be regarded correct to one figure only.

It is interesting to compare this result with the expression, 22×10^{18} ergs per second, recently obtained by H. Jeffreys[*] for the mean rate of energy dissipation in the semi-diurnal tidal motion of all the shallow seas of the world. The method there used is due to G. I. Taylor,[†] and is based on the assumption that the dissipation is caused by direct friction of moving water on the bed of the sea, the velocity of slip being taken equal to the surface current velocity. The rate of energy dissipation is assumed to be expressible in the form[‡] $k\rho v^2$ per unit area of bottom, ρ being the density, v the velocity of slip, and k a constant which is taken as equal to that found experimentally for the wind blowing over grass land.[§]

The Relations between Currents, Salinities, and Temperatures.

An estimate of the magnitude of ocean currents can be obtained by simple methods from a knowledge of the distributions of the salinity or of the temperature of the sea water.

Let s denote the salinity at any point at time t, that is, the mass of salt dissolved in unit mass of the water which at that instant surrounds the point considered. Making the usual assumption that each small portion of water retains its salinity, we have the equation

$$\frac{\mathrm{D}s}{\mathrm{D}t} = 0,$$

---

[*] 'Phil. Trans.,' A, vol. 221, p. 239 (1920).

[†] 'Phil. Trans.,' A, vol. 220, p. 1 (1919).

[‡] Note that this is not of the same form as F or F′ in equations (7) and (8).

[§] There seems to be no assurance that the agreement between this somewhat arbitrary determination of k and another estimate of its value obtained from an extrapolation of Bazin's formula for the flow of water in a river may not be largely accidental. The various hydraulic formulæ for such motion seem to be empirical; and others, which agree equally well with the observations, when extrapolated give results of entirely different orders of magnitude. *Cf.* Lord Rayleigh, 'Sci. Papers,' vol. 6, p. 602.

where the differentiation follows the path of the water particle. If this equation is applied to a particle at the surface of the sea, it can be written in the form

$$\frac{\partial s}{\partial t}+\mathrm{Q}s'\cos\chi+w\frac{\partial s}{\partial z}=0, \tag{11}$$

where Q is the horizontal and w the vertical velocity of the water, the axis of z being measured upwards, s' is the maximum surface gradient of s at time t, and χ is the angle between the directions of Q and s'.

Similarly, if ρ is the density and T the temperature of the water at the point at time t, the assumption that the conduction of heat is negligible leads to the equation

$$\mathrm{D}(\rho\mathrm{T})/\mathrm{D}t=0,$$

while the equation of conservation of mass is

$$\mathrm{D}\rho/\mathrm{D}t=0.$$

From these we deduce

$$\mathrm{DT}/\mathrm{D}t=0,$$

which as before leads to a result of the form

$$\frac{\partial\mathrm{T}}{\partial t}+\mathrm{QT}'\cos\psi+w\frac{\partial\mathrm{T}}{\partial z}=0, \tag{12}$$

where T′ is the maximum surface gradient of T at time t, and ψ is the angle between the directions of Q and T′.

The value of Q could be found from either of the equations (11), (12), if the necessary related sets of observations of the salinity or of the temperature were available. At present, it is possible only to indicate the orders of magnitude of the several terms in the equations. It is probable that each of the three variables (current, salinity, temperature), is the sum of a number of periodic functions of the time, together with a part independent of t. The equations should therefore be applied to each of these component parts; but, in view of the scanty nature of the data, this is quite impracticable.

The equations have been applied numerically to the North Atlantic, taken as a whole, rather than to particular points in it, the aim, as stated above, being to obtain orders of magnitude only, and not definite numerical results. A very large number of observations of salinity and of temperature, spread over a period of two years, have been collected by H. N. Dickson*; and, from these, he has constructed charts showing the isohalines and the isotherms for each of the twenty-four months of the period covered. Their most striking feature is the difference between charts relating to the same month of two consecutive years. It is probable that this variation can be attributed to meteorological causes; for it is clear that such events as the breaking up of the polar ice will have a profound effect on the chemical and physical conditions in the Ocean.

---

* 'Phil. Trans.,' A, vol. 196, p. 66 (1901).

Adopting the usual practice of measuring the salinity of the water by the number of parts by weight of salts dissolved in a thousand parts of water, we find as rough mean values*:—

$$s' \,(=)\, 0\cdot2 \text{ per 100 sea miles,}$$

$$\mathrm{T}' \,(=)\, 2^\circ \text{ C. per 100 sea miles,}$$

or using C.G.S. units,

$$s' \,(=)\, 3 \times 10^{-10}\, s, \qquad \mathrm{T}' \,(=)\, 2 \times 10^{-7}. \tag{13}$$

In order to obtain the time-gradients of the salinity from the charts, three representative points in latitude 50° N., with longitude 20°, 30°, 40° W., were taken, and for each the salinity-time curve was constructed by interpolation. The graphs naturally show some irregularities; but it is evident that the predominant parts have a period of six months, their amplitudes being about 0·1 for the first two stations, and 0·4 for the third. Taking 0·2 as the mean amplitude, we find for this term

$$\partial s/\partial t \,(=)\, 2 \times 10^{-9}\, s. \tag{14}$$

The effect of the very small daily variation in salinity is small compared with this, notwithstanding the more rapid rate of variation.

In a similar manner, the time-gradients of the temperature can be obtained. As could have been foreseen, the predominating term has a period of a year, its amplitude being about 3° C. in each case. This is approximately half the mean of the range of surface temperature of all the oceans. There is, in addition, a daily component, whose amplitude is about 0·3° C. This is, therefore, the more important part, and we may suppose

$$\partial \mathrm{T}/\partial t \,(=)\, 10^{-5}. \tag{15}$$

The vertical gradients of the salinity and the temperature are less easy to obtain. Thoulet† has calculated that the mean value of the salinity of the North Atlantic for depths less than 50 m. is 35·82, and for depths from 50 m. to 100 m. is 35·99. Using these figures, we find $\partial s/\partial z \,(=)\, 3 \times 10^{-5}$. Other observations confirm the accuracy of this estimate. Thus we may take

$$\partial s/\partial z \,(=)\, 10^{-6}\, s, \tag{16}$$

---

* The symbol (=) is used to denote equality of orders of magnitude.

† 'Résultats des Campagnes scientif. du Prince de Monaco,' vol. 29, p. 95.

while from a consideration of the vertical variation of temperature at many stations we infer

$$\partial T/\partial z \, (=) \, 10^{-4}. \tag{17}$$

The most important part of the vertical velocity of the water particles at the surface of the ocean at points far removed from land may be supposed due to a semi-diurnal wave of amplitude of the order 30 cm. Hence

$$w \, (=) \, 5 \times 10^{-3}. \tag{18}$$

From equations (13), (14), (16), (18) we see that the terms of (11) are of orders

$$2 \times 10^{-9} s, \qquad 3Q \times 10^{-10} s, \qquad 5 \times 10^{-9} s,$$

which are in the ratio

$$20, \qquad 3Q, \qquad 50.$$

Thus Q is of the order 20 cm. per sec. or, say, half a knot.

In a similar manner we see from (13), (15), (17), (18) that the terms of (12) are of orders

$$10^{-5}, \qquad 2Q \times 10^{-7}, \qquad 5 \times 10^{-7},$$

which are in the ratio

$$100, \qquad 2Q, \qquad 5.$$

Thus Q is of the order 50 cm. per sec. or one knot.

The mean current velocity of the surface of the Atlantic Ocean is rather less than a knot. The agreement of these two rough estimates with the observed result is satisfactory, and seems to confirm the mutual consistency of the data employed.

The Constitution of the Alloys of Aluminium, Copper, and Zinc containing High Percentages of Zinc.

By JOHN L. HAUGHTON, D.Sc., F.Inst.P., and KATHLEEN E. BINGHAM, M.Sc.
(both of the National Physical Laboratory).

(Communicated by Dr. W. Rosenhain, F.R.S. Received May 5, 1920.)

[PLATES 1 AND 2.]

The following paper constitutes one of a series of researches on alloys, particularly on those containing aluminium, which is being carried out at the National Physical Laboratory, under the supervision and guidance of Dr. W. Rosenhain, F.R.S.

Introduction.

The ternary system copper-aluminium-zinc is of considerable interest, alike from the practical as from the theoretical point of view.

Portions of the ternary system have already been worked out. The copper-rich alloys have been investigated by Carpenter and Edwards.* The liquidus surface has been determined by these authors for the area CuBC (fig. 1), and the complete constitution for the area CuAC. As yet no constitutional diagram has been published for the aluminium-rich alloys of the series, but this forms the subject of one of the collateral researches being carried out in the National Physical Laboratory. The zinc corner (that part marked ZnDEF, fig. 1) is the subject of the present research.

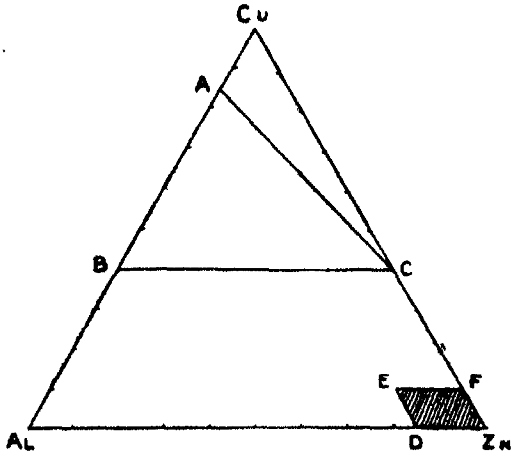


FIG. 1

---

* H. C. H. Carpenter and C. A. Edwards, 'Internat. Zeits. f. Metallographie,' vol. 2, Part 4, p. 209 (1912).

The Binary Systems.

1.—*Copper-Aluminium.*—The system copper-aluminium has been investigated by Curry,[*] Gwyer,[†] Carpenter and Edwards,[‡] and others.

The equilibrium diagram may be said to be fairly well established in the main, although there are certain details which require to be further investigated. As far as the present research is concerned, however, this diagram is of little importance.

2. *Aluminium-Zinc.*—The alloys of aluminium and zinc have been investigated by Shepherd,[§] Rosenhain and Archbutt,[||] and Bauer and Vogel.[¶]

3. *Copper-Zinc.*—That part of the copper-zinc diagram with which the present work is concerned has been investigated by Shepherd and Blough,[**] by Tafel,[††] and others. Both diagrams are manifestly incomplete. It was considered advisable, therefore, as a preliminary to the main investigation, to redetermine that part of the equilibrium diagram of copper-zinc alloys containing from 90 per cent. to 100 per cent. of zinc.

Before referring to the results obtained in the present research, it will be advisable to describe in some detail the methods and apparatus employed in the investigation. The methods were:—

1. Thermal curves.
2. Microscopic examination of annealed and quenched specimens.
3. Electrical conductivity measurements.

1. *Thermal Curves.*—The thermal curves were taken in a "gradient" furnace of the type described by Rosenhain in his paper on "Some Appliances for Metallographic Research."[‡‡] As this particular furnace was not required for high temperature work, the tube carrying the winding was made of iron, in order to obtain a more gradual temperature gradient than was possible by the use of a fireclay tube. The specimens employed weighed about 50 grm. and were carried in a small clay pot, which was made with a tube of the same material passing in from one side. The pot was carried in an iron cylinder, which was hung from the winding gear by three "Nichrome" wires,

---

[*] H. E. Curry, 'Jnl. Phys. Chem.,' vol. 11, p. 425 (1907).

[†] A. G. C. Gwyer, 'Zeits. f. Anorg. Chem.,' vol. 57, p. 113 (1908).

[‡] H. C. H. Carpenter and C. A. Edwards, 'Proc. Inst. Mech. Eng.,' vol. 57 (1907).

[§] E. S. Shepherd, 'Jnl. Phys. Chem.,' vol. 9, p. 504 (1905).

[||] W. Rosenhain and S. L. Archbutt, 'Phil. Trans.,' A, vol. 211, p. 315 (1911).

[¶] O. Bauer and O. Vogel, 'Internat. Zeits. f. Metallographie,' vol. 8, p. 107 (1916).

[**] E. S. Shepherd and Blough, 'Jnl. Phys. Chem.,' 1904.

[††] V. E. Tafel, 'Metallurgie,' vol. 5, p. 375 (1908).

[‡‡] W. Rosenhain, "Some Appliances for Metallographic Research," 'Journ. Inst. of Metals,' vol. 13, No. 1, p. 160 (1915).

passing through holes in a ring, which kept them in place. The suspending wire and the two thermocouple wires passed through a silica tube, and the junction of the couple was inserted into the tube in the side of the crucible, the orifice of which was then packed with asbestos wool. This tube was made very thin at the end, so as to reduce the thermal lag to a minimum. The couple employed was an iron-constantan one, which was carefully calibrated before and after the curves were taken, the melting points of tin, lead, and zinc being used for the purpose. The two sets of calibrations show only negligible differences.

The specimen was raised and lowered in the furnace by means of a weight, the speed of the motion being controlled by the movement of a piston in a cylinder of oil. Readings were taken every 2° C., and the rate of heating or cooling was of the order of 5° C. per minute.

The curves were recorded on the Rosenhain plotting chronograph,* the paper being reversed on the drum between heating and cooling, so that the temperature ordinates were the same for the two curves.

2. *Microscopic Examination.*—The microscopic examination calls for no special comment. The specimens were annealed and quenched before examination, except in special cases where very slowly cooled or chill cast alloys were examined. When it was necessary to anneal for very long periods, this was carried out in a Haughton-Hanson thermostat,† the specimens being tied together and to the thermo-couple with a thin "Nichrome" wire. In order to avoid the risk of contamination, the specimens were insulated from the wire and from each other by a thin strip of asbestos paper. When the annealings were completed, the specimens were rapidly withdrawn from the furnace by means of the thermo-couple and plunged into a basin of cold water. In these cases where rapidity of quenching was of more importance than prolonged annealing, *e.g.*, in the case of the determination of the solidus, a Rosenhain quenching furnace was used. As, however, it has been shown by Thorneycroft and Turner,‡ that, *in vacuo*, copper-zinc alloys containing about 10 per cent. of copper lose from 20 per cent. to 25 per cent. of their zinc at temperatures similar to those at which the annealings were carried out, it was not considered advisable to use a vacuum in order to obtain the desired inrush of quenching water. Instead of this, a rapid stream of water from a tap was injected into the furnace, and this washed the

---

* W. Rosenhain, *loc. cit.*

† J. L. Haughton and D. Hanson, "A Thermostat for Moderate and High Temperatures," 'Journ. Inst. of Metals,' vol. 14, p. 145 (1915). "Further Notes on a High Temperature Thermostat," *ibid.*, vol. 18 (1917).

‡ W. E. Thorneycroft and T. Turner, "Behaviour of Copper-Zinc Alloys when Heated in a Vacuum," *ibid.*, vol. 12, p. 214 (1914).

specimen into a vessel of water placed at the end of the furnace. A very rapid quench was obtained by this means.

In a few cases it was necessary to cool the alloys slowly, from a point above the liquidus, to a temperature intermediate between this line and the solidus. In order to do this, the specimens were placed in a small glass crucible, the thermo-couple being inserted into a narrow glass tube sealed into it. Fragments of the alloys were placed in this crucible, which was then hung, by means of very thin copper wire, inside a vertical tube furnace, mounted over a beaker of water. The furnace was run in parallel with a thermostat, fitted with the slow cooling device described in the paper referred to above, and the desired rate of cooling was obtained. When the specimen had reached the temperature at which it was to be quenched, the supporting wire was fused electrically, and the pot and specimen fell into the water.

It was found possible, in many cases, to etch the alloys by exposing them to the atmosphere, but this was not a suitable method for detecting the first appearance of the ϵ constituent. Numerous reagents were tried, and a very satisfactory one was found to be that of Tinoféef.* This reagent was suitable for all the alloys investigated, and was almost invariably used. In a few cases, however, specimens were given a preliminary air etch, and then treated with Tinoféef's reagent, while in other cases electrolytic etching with 4 per cent. hydrochloric acid in water was used.

Towards the end of the research, it was found that, to obtain good contrast between the three constituents, the most satisfactory method was to use an electrolytic etch in a solution consisting of twenty drops of concentrated hydrochloric acid in 50 c.c. of alcohol.

3. *Electrical Conductivity.*—The electrical conductivity of the alloys was determined by measuring the drop of potential along a measured length of the specimen, through which a constant current was maintained. The apparatus employed, and the method of carrying out the tests, have been described by one of the authors in a paper read before the Faraday Society.†

The Equilibrium Prism.

It is well known that the equilibrium conditions of alloys of three metals can be represented by a triangular prism. Sections through such a prism, taken parallel to one of the sides, represent a series of alloys in which one of the metals is present in constant quantity, and the other two are variable, while sections parallel to the base represent the entire series at a definite temperature.

---

* Tinoféef's reagent is made by dissolving 6 grm. of chromium oxide in 94 grm. of nitric acid. A few drops of this solution are then added to 50 c.c. of water.

† J. L. Haughton, "The Measurement of Electrical Conductivity in Metals and Alloys at High Temperatures," 'Trans. Faraday Society,' vol. 16 (1920).

Figs. 2–9 represent a series of such sections parallel to the aluminium-zinc side of the prism, the copper contents being 0, 1, 2, 3, 4, 5, 7, and 9 per cent. respectively. Fig. 2 is copied from Rosenhain and Archbutt's diagram, referred to above. Figs. 10–17 are sections parallel to the copper-zinc side of the prism, the aluminium contents being 0, 2, 4, 6, 8, 10, 12, and 15 per cent. Fig. 18 is a contour plan of the liquidus surface, while figs. 19, 20 and 21 are sections cut through the complete series at temperatures of 400° C., 300° C., and 200° C.

From such a series of cross-sections, a temperature-concentration model can be constructed.

At Dr. Rosenhain's suggestion, the authors constructed a wire model,* and figs. 22, 23, and 24 have been drawn from this model, the various planes being cross-hatched, so as to represent progressive cooling of the alloys. These figures, together with the various sections, are explained below, under their proper headings.

The liquidus curves of the various sections have been determined solely from the thermal curves, except in one case, where the results obtained from the temperature-conductivity curve were also used. In some cases, the indications on the heating curves of the completion of fusion were so incomplete that it was considered advisable to neglect them and use the data from cooling curves only.

The various solidus curves were obtained from the thermal analysis when the solidus happened to be an eutectic or a transformation, otherwise they were obtained from the examination of quenched specimens, and in one case confirmed by electrical conductivity measurements. Reactions occurring in the solid were detected by the thermal curves and confirmed by electrical conductivity methods, and by microscopic examination of annealed and quenched specimens.

Limits of solubility were determined by microscopic examination of specimens which were quenched after long annealings in the thermostat furnace.

Nomenclature.—In order to avoid referring to the composition of the alloys in detail, a nomenclature has been adopted, in which numerals representing the approximate copper and aluminium contents are separated by the letter Z, which represents zinc. Thus, 5Z7 represents an alloy whose composition is approximately :—

---

* W. Rosenhain, "Model representing Ternary Alloys," 'Journ. Inst. of Metals,' vol. 23, p. 247 (1920).

	Per cent.
Cu	5
Zn	88
Al	7

As the alloys were very carefully made up from previously made and

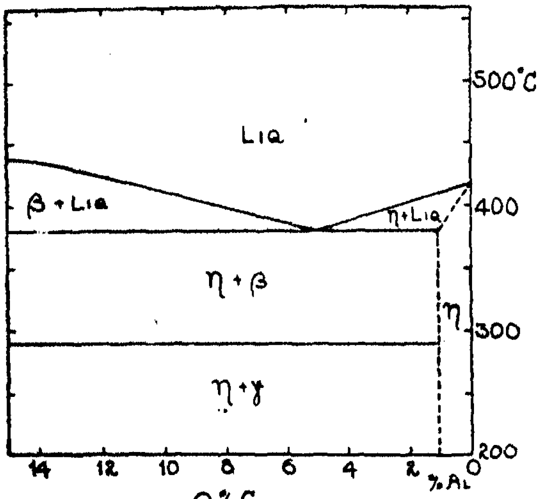


FIG. 2.

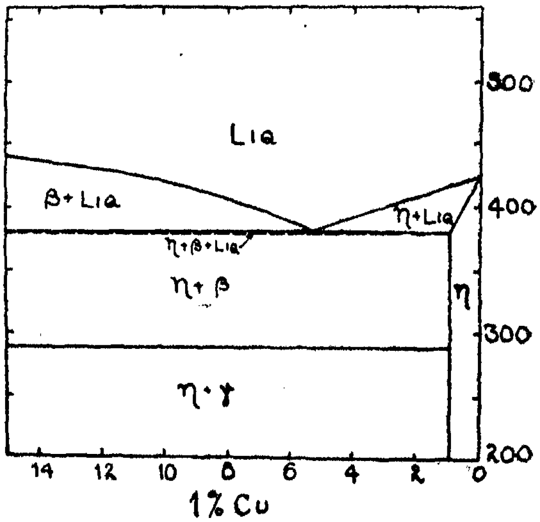


FIG. 3.

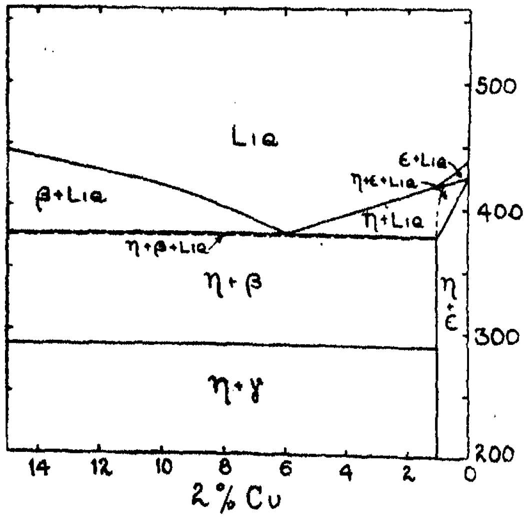


FIG. 4.

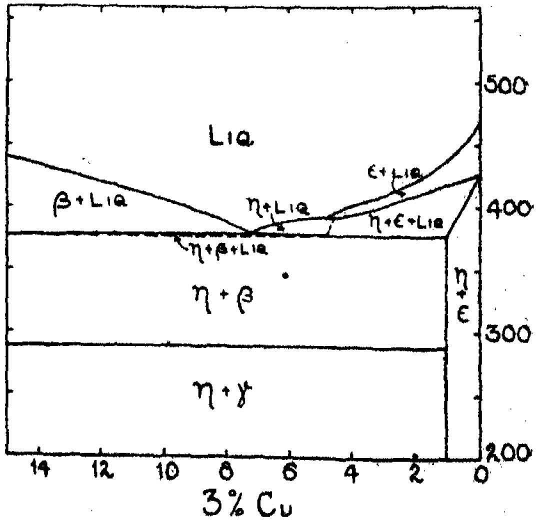


FIG. 5.

analysed "basis" alloys (50 per cent. Al, 50 per cent. Cu, and 90 per cent. Zn, 10 per cent. Cu respectively), the actual composition varies but slightly from that indicated by the nomenclature.

The Copper-Zinc Series.—As has been indicated earlier in this paper, the first step to be taken was the redetermination of the constitution of the copper-zinc alloys containing between 90 and 100 per cent. of zinc. The

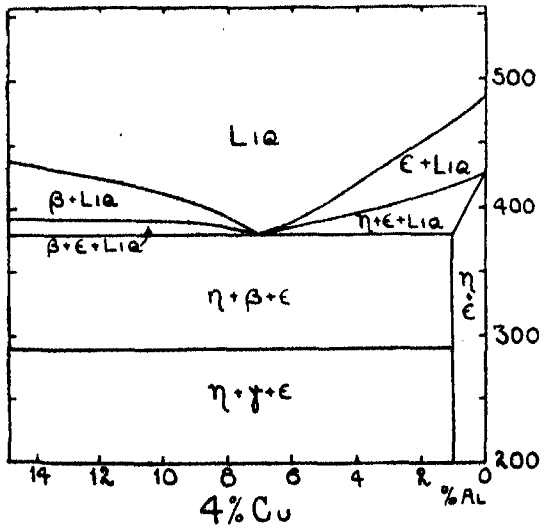


FIG. 6.

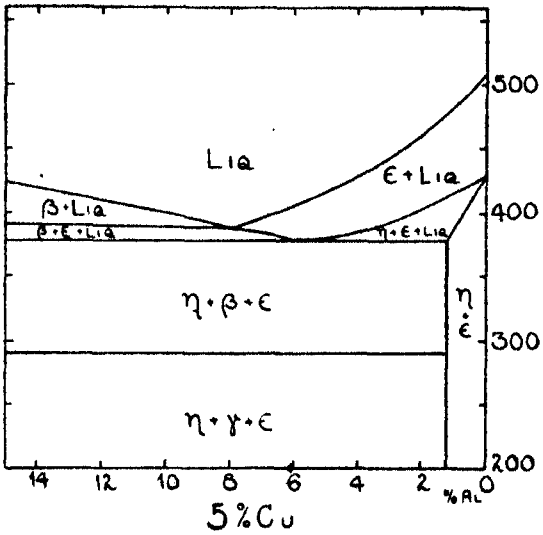


FIG. 7.

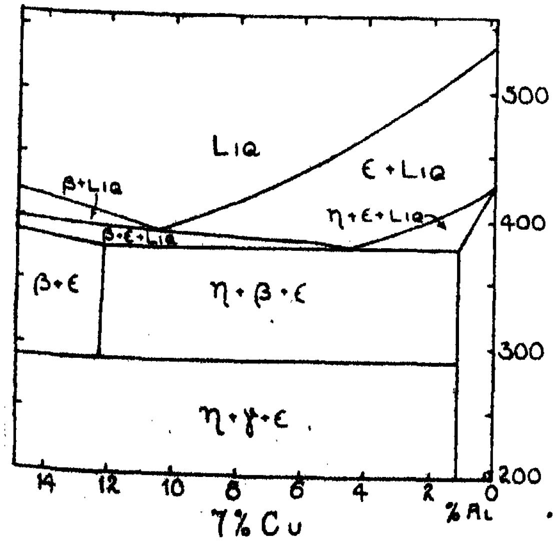


FIG. 8.

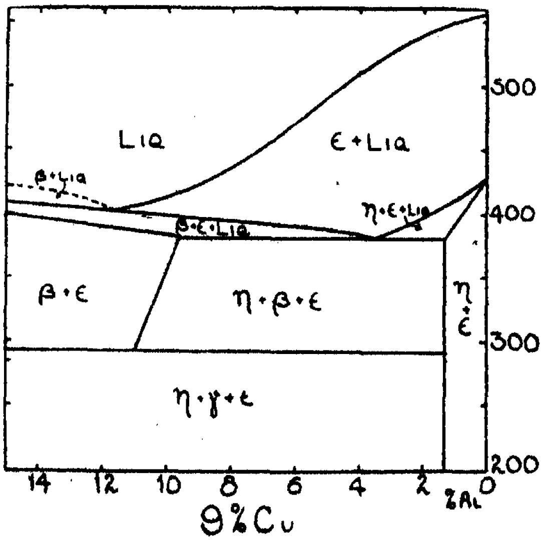


FIG. 9.

result obtained by the authors is given in fig. 10. The liquidus agrees well with that of Tafel, while the temperature of the reaction

$$\epsilon + \text{liquid} = \eta$$

is found to be 426° C. This, again, is in excellent agreement with Tafel, who gives a temperature of about 425° C. for the reaction as compared with 419° C., which is the temperature indicated by Shepherd and Blough.

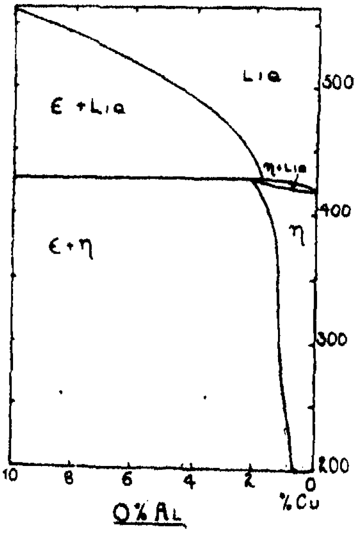


Fig. 10.

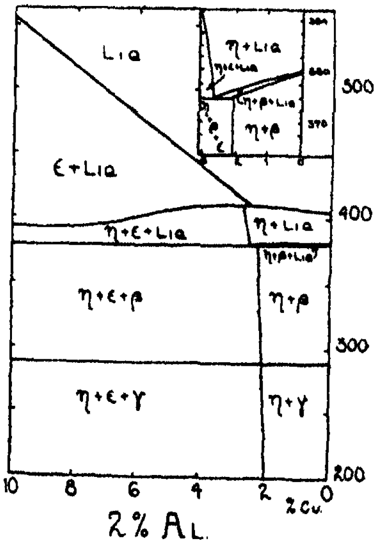


Fig. 11.

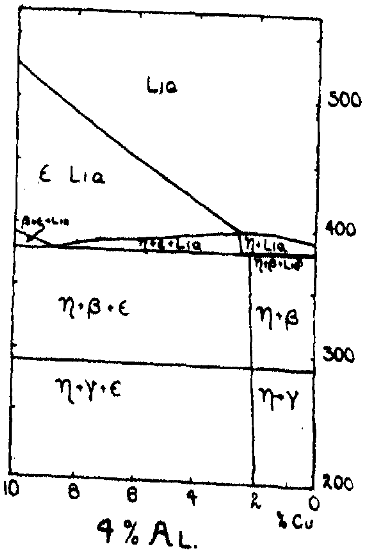


Fig. 12.

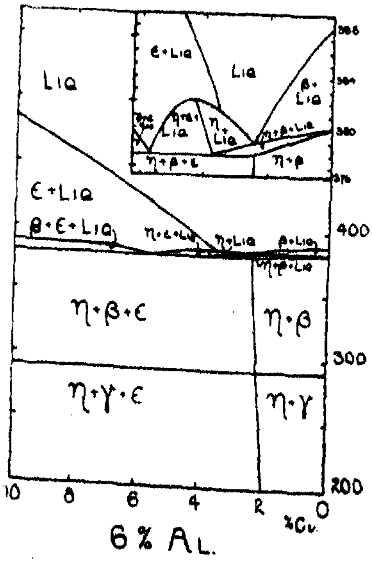


Fig. 13.

The determination of the solidus between the limits of 0 and 2 per cent. of copper was a matter of considerable difficulty, as the maximum range between the liquidus and solidus was found to be 5° C. By quenching the specimens at temperatures rising by about 1° C. at a time, and examining them under the microscope, it was possible to detect the first appearance of fusion spots, *i.e.*, places which were molten at the moment of quenching. The alloys which were examined in this way were ¼Z0, 1Z0, 1½Z0, and 2Z0.* By this means it was found possible to determine the solidus with an

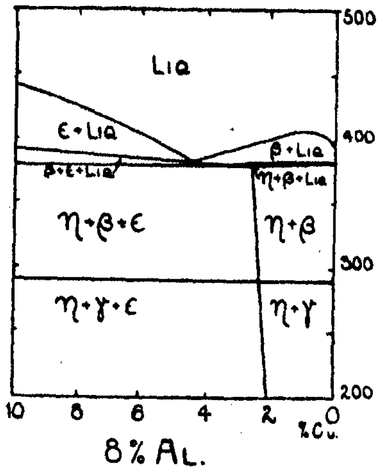


Fig. 14.

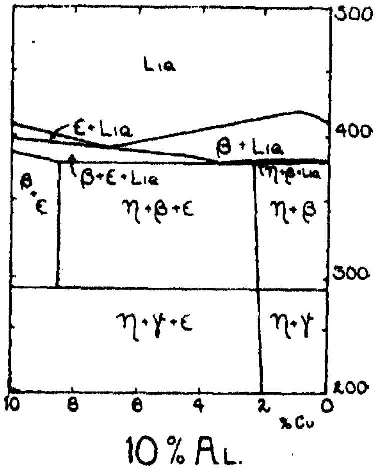


Fig. 15.

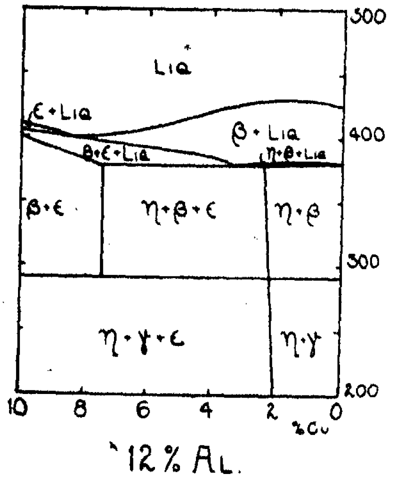


Fig. 16.

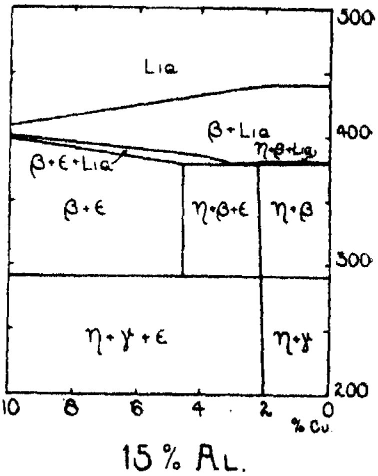


Fig. 17.

---

* Analysis showed that the composition of these alloys was 0·26 per cent., 0·95 per cent., 1·30 per cent. and 1·80 per cent. of copper respectively.

error probably not greater than ± 1° C. Confirmation of this was found from a temperature-electrical conductivity curve taken of alloy 1Z0, which is reproduced in A, fig. 25. The more or less horizontal part of this curve

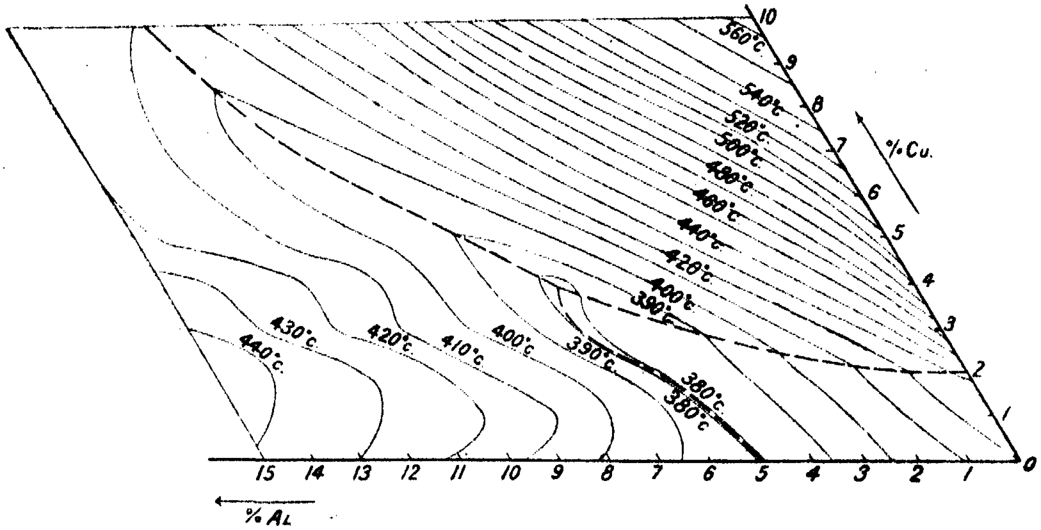


FIG. 18.

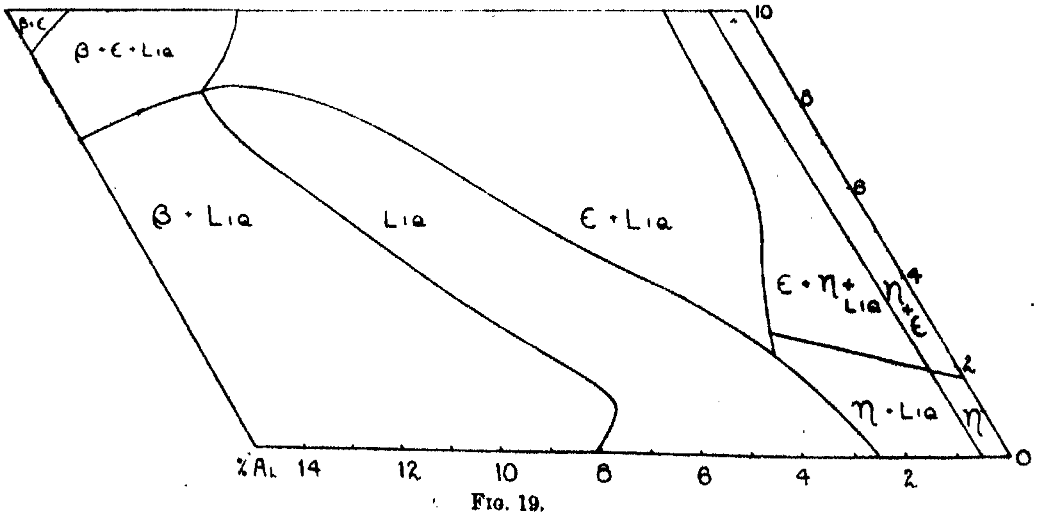


FIG. 19.

represents the resistance of the material during the change from the liquid to the solid state, and will be seen to have a range of about 5° C. For purposes of comparison, the curve for pure zinc is given as B in fig. 25.

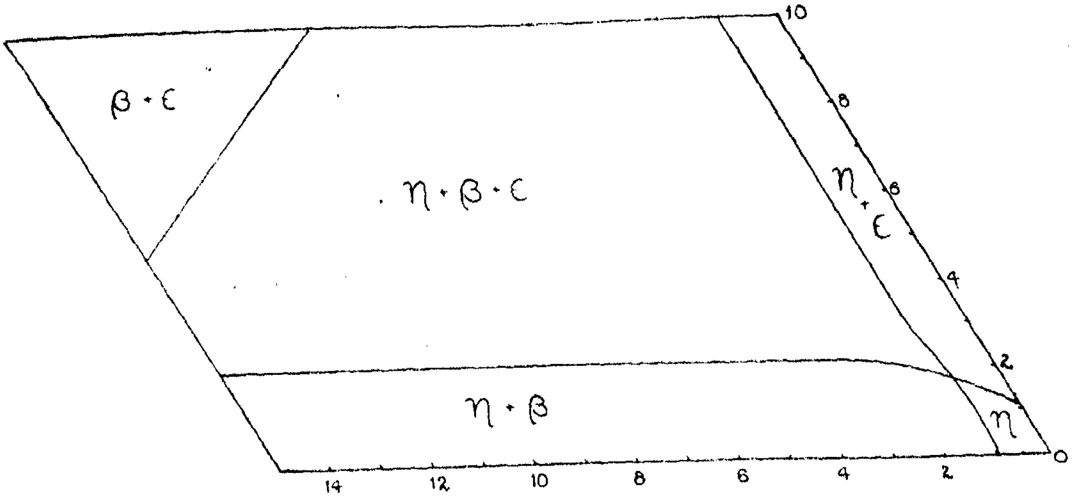


Fig. 20.

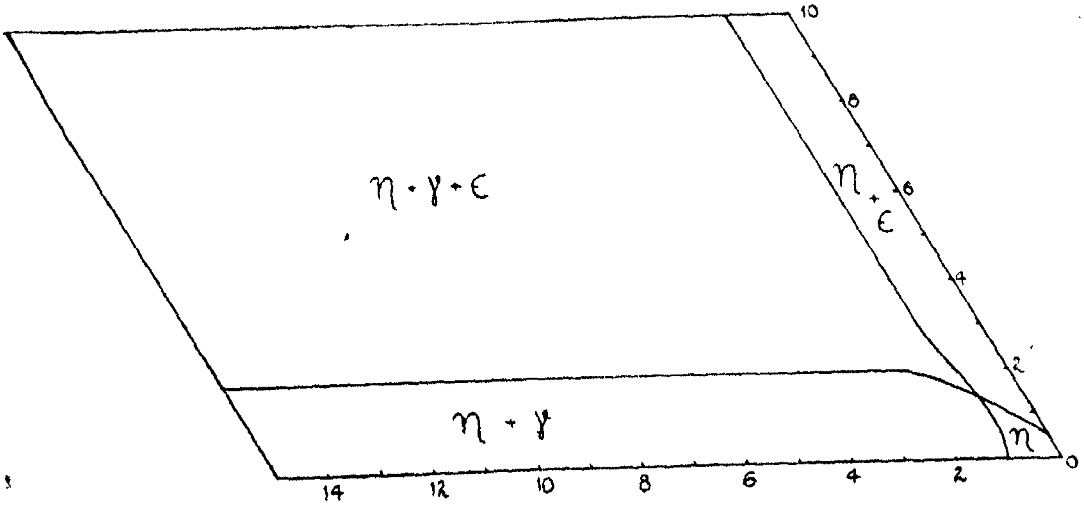


Fig. 21.

The Ternary Diagram.

The results obtained in the investigation of the ternary system are shown in figs. 3 to 17.* The case of the alloys containing 1 per cent. copper is very simple, and the same may be said of those having a copper-content of 2 per cent., with the exception of a small corner near the pure copper-zinc end (figs. 3

---

* The α constituent of the Al-zinc series, and the η constituent of the copper-zinc series are isomorphous. In order to avoid confusion, the term α has been omitted altogether, and the homogeneous zinc rich solid solution of either copper or aluminium or both, has been called η in this paper.

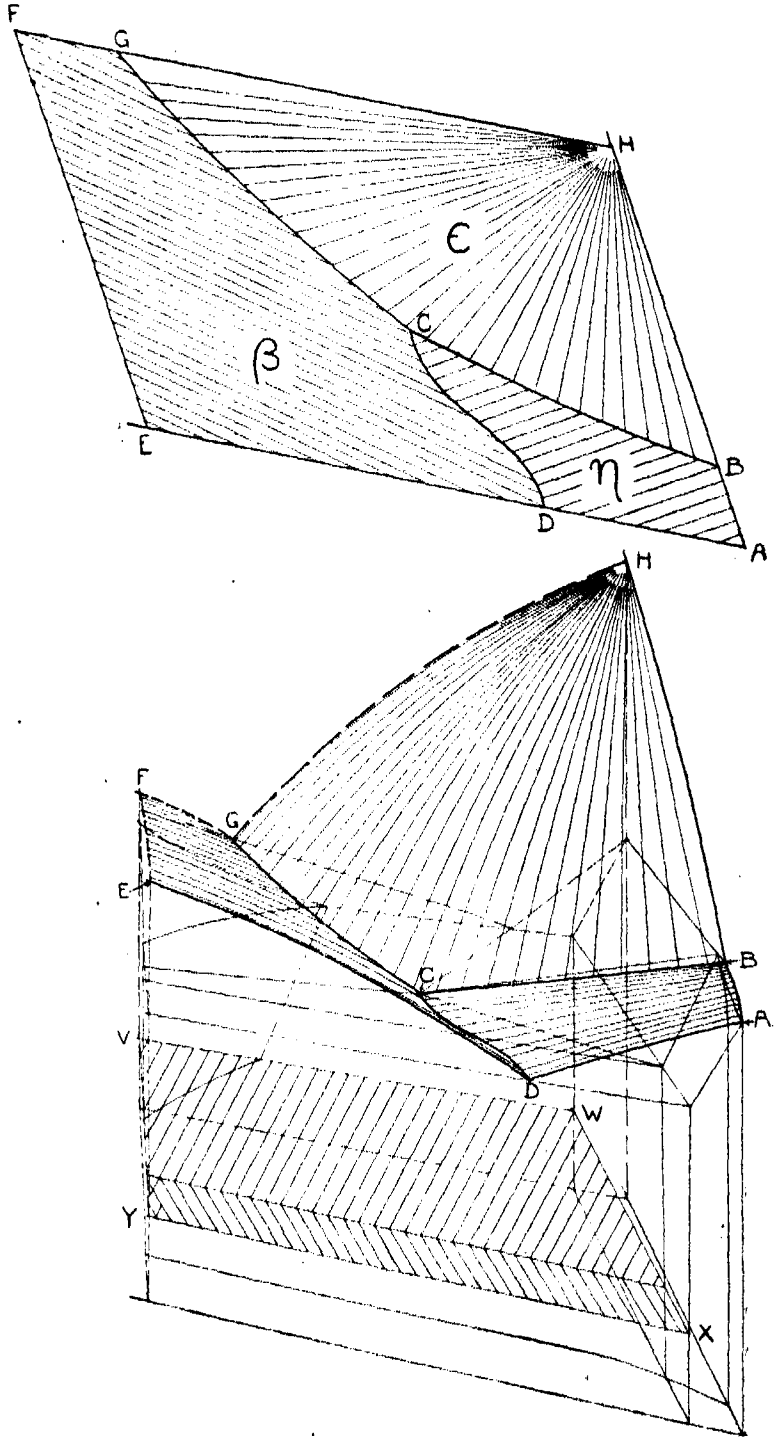


Fig. 22.

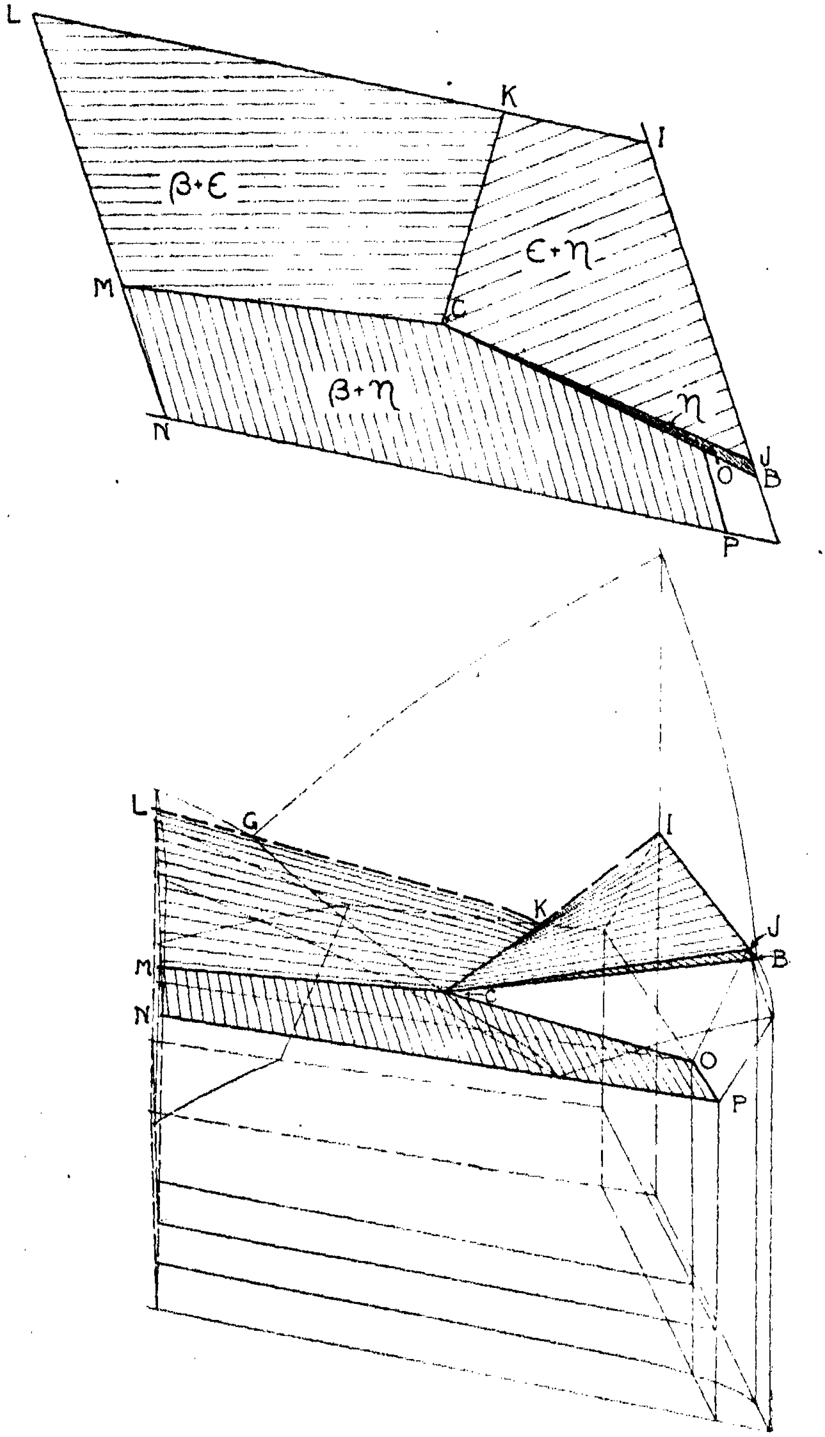


Fig. 23.

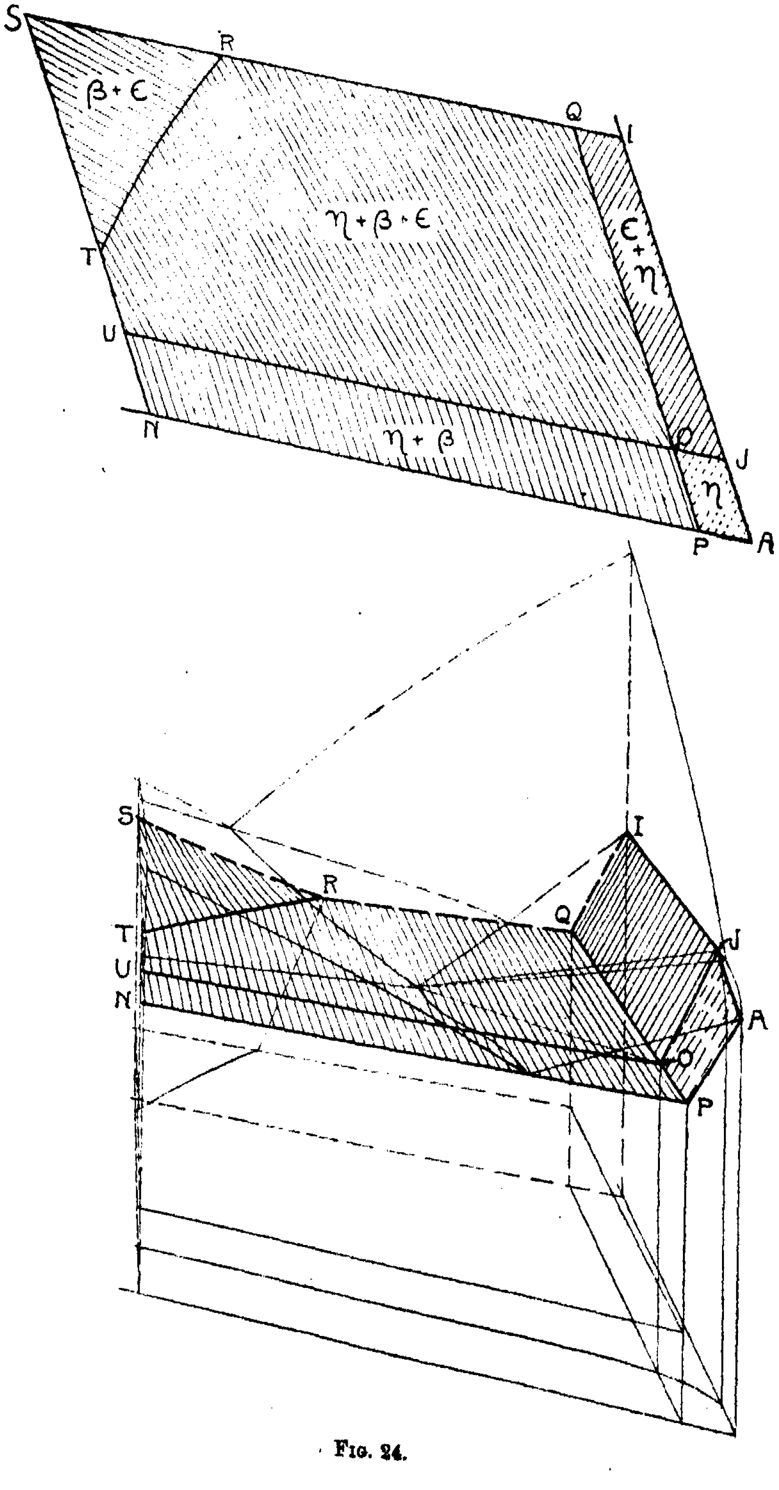


Fig. 24.

and 4). These diagrams, with the exception just referred to are, practically, repetitions of the aluminium-zinc diagram as far as they have been investigated (*i.e.*, up to 15 per cent. Al). The eutectic of the series, however, being a binary complex in a ternary system, cannot solidify at a constant temperature, but the solidification must take place over a range which must be lower than the temperature at which the pure eutectic forms. The range appears, however, to be so small as to be beyond the limits which can be detected by the pyrometer, and one of the lines has therefore been shown dotted on the diagram. The evidence for the lowering of the temperature seems somewhat more marked, though it is still very slight, amounting to about 2° C. between the temperature of solidification of the pure binary eutectic and that of the ternary eutectic which occurs in the alloys containing 3 per cent. of copper and over.

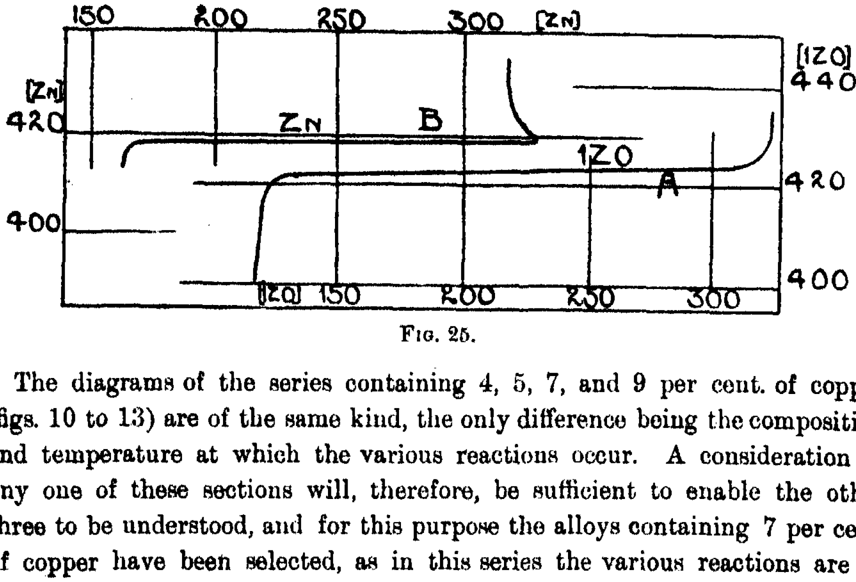


Fig. 25.

The diagrams of the series containing 4, 5, 7, and 9 per cent. of copper (figs. 10 to 13) are of the same kind, the only difference being the composition and temperature at which the various reactions occur. A consideration of any one of these sections will, therefore, be sufficient to enable the other three to be understood, and for this purpose the alloys containing 7 per cent. of copper have been selected, as in this series the various reactions are all well developed.

On cooling an alloy containing 7 per cent. of copper and 0·5 per cent. of aluminium, the first substance to separate from the liquid is the copper-zinc solid solution ϵ (containing a small amount of aluminium in solution, as is shown by the depression of its freezing point). This separation commences at a temperature of 523° C. and continues until 417° C. is reached, when the solid reacts with the liquid according to the equation

$$\epsilon + \text{liquid} = \eta.$$

In the pure copper-zinc series this reaction takes place at a constant temperature, but in the ternary system, not only is the temperature depressed, but the reaction continues over a range: in the case of the alloy under consideration, until 390° C. is reached, when the whole of the liquid is used up, and ϵ and η are left. This is shown in fig. 26 (Plate 1). In this and in figs. 27, 28, and 29 are shown photomicrographs of alloys which have been cooled slowly from a temperature considerably above the liquidus down to 350° C.

Considering next the alloy with the same copper-content, but with 2 per cent. of aluminium, we find as before, solidification commences by the separation of the ϵ crystals from the liquid, the temperature in this case being 495° C. At 398° C. the reaction giving rise to η commences, continuing with falling temperature until 378° C. is reached, when the remaining liquid solidifies at a constant temperature in the form of a ternary eutectic (fig. 27). The nature and manner of formation of this eutectic will be discussed later.

An alloy containing 8 per cent. of aluminium (7Z8) behaves very differently. As in the case of 7Z½ and 7Z2 the first constituent to separate is ϵ, the separation commencing at 415° C. This, however, never reacts with the liquid to form η, but when the temperature has fallen to 388° C. a binary complex of ϵ with the β solution of the aluminium-zinc series (Al_2Zn_3) commences to separate from the liquid. This separation proceeds until the temperature of 378° C. is reached, when the ternary eutectic solidifies. The alloy, when solid, will therefore show primary separation of ϵ and sub-primary or secondary separation of the β-complex in a matrix of the ternary eutectic (fig. 28).

Referring next to the alloy containing 10·5 per cent. of aluminium (7Z10½). Here the first solid to separate is no longer ϵ but the $\beta-\epsilon$ complex referred to above. This separation commences at 392° C. and continues until the ternary eutectic solidifies. The solid alloy will therefore consist of primary spherulitic dendrites of the complex embedded in the ternary eutectic (fig. 29).

The first separation from the alloy 7Z12 takes place at a temperature of 402° C., and consists of β-crystals. These continue to separate over a range of 8° C., when the complex begins to solidify, and at 378° C. as before, the remaining liquid freezes as the ternary eutectic. The just solid alloy will, therefore, consist of primary β and secondary $\beta-\epsilon$ complex with a certain amount of ternary eutectic (see fig. 30). Finally, the alloy 7Z15 commences to solidify in a similar manner at 420° C., and at 400° C. the $\beta-\epsilon$ complex begins to separate, this separation continuing until the temperature of 390° C. is reached, when the whole alloy is solid (fig. 31). It will be seen that in this case there will be no ternary eutectic present. Figs. 30 and 31 are photographs of alloys which have been chill cast and annealed for long periods to remove the last traces of metastable ternary eutectic. This accounts for the difference in the scale of the structure compared with that of figs. 27, 28 and 29.

The nature of the $\beta-\epsilon$ complex is particularly interesting. It is the only case with which the authors are acquainted where a binary complex is formed in a ternary system without any corresponding eutectic existing in any of the binary systems. In this case, one of the constituents of the complex belongs essentially to one of the binary systems, and the other to the other system. The constituents are β, which is the compound Al_2Zn_3, probably containing a certain amount of copper in solution, and ϵ, which is probably a solution of zinc in the compound $CuZn_6$.

In order to show the method of formation of this eutectoidal complex, the alloy 9Z12 was slowly cooled from above the liquid, and quenched at 381° C. It was desired to obtain an alloy in which the primary separation consisted solely of the complex, and which also contained little or no ternary eutectic. It will be seen that 9Z12 is the alloy which comes nearest to this specification. Fig. 32 shows the microstructure of this alloy when slowly cooled from above the liquid, and quenched at 381° C. It will be seen that, as was to be expected, there is only a primary separation of the $\beta-\epsilon$ complex, together with a small quantity of chilled liquid.

In the alloys containing 3 per cent. of copper, no thermal evidence can be obtained for the existence of the binary complex, although the microstructure shows that it is present. The reason of this doubtless is that in this section the separation of the aluminium-zinc eutectic only just precedes that of the ternary eutectic, and the two reactions occur at so nearly the same temperature that they are not separated by the thermo-couple. In sections of the model representing alloys containing 4 per cent. of copper and upwards it is easily detected.

The ternary eutectic must now be discussed. The three phases composing it are obviously η, β, and ϵ, but the manner in which it is formed is not always the same. In the majority of cases it is formed as though it were a binary eutectic, one of whose constituents is the $\beta-\epsilon$ complex, and the other is η, which arises from the reaction between primary ϵ and the liquid. This can be seen well in fig. 37.

While the eutectic is generally formed as outlined above, this is not always the case. Over a small range of composition, between about 5 and 15 per cent. of aluminium, the copper-content varying between 2·5 and 4 per cent., the eutectic resembles in its method of formation a binary eutectic composed of the $\beta-\epsilon$ complex, and the complex consisting of the depressed $\eta-\beta$ eutectic, of the aluminium-zinc series. In a third case, the ternary eutectic is formed from this $\eta-\beta$ complex and ϵ.

The exact composition of the ternary eutectic has been determined mainly from the microstructure. A very small change in the composition makes a very great increase in the amount of eutectic present. This will be seen by comparing fig. 33 (alloy 3Z6) with fig. 35 (alloy 4Z6). In the former, there is only a small trace of ternary eutectic in the boundaries of the crystals; in the latter, a very large part of the alloy consists of this eutectic, the change being brought about by the addition of only 1 per cent. of copper. Similarly, considering figs. 35, 36, and 37 (alloys 4Z6, 4Z7, and 4Z8), in fig. 35 there are large crystals of primary ϵ and some evidence for primary η as well. Fig. 36 represents an alloy which must be very near the ternary eutectic composition, as there is no constituent present which can definitely be said to be a primary separation, while in fig. 37 it is fairly evident that the $\beta-\epsilon$ complex is the primary separation. It would appear from this that the ternary eutectic cannot be very far removed from the composition:—Cu 4 per cent., Al 7 per cent., Zn 89 per cent.

The occurrence of the ternary eutectic appears to be limited to a very small part of the system, and it is obvious that it is not the same eutectic as that referred to by Carpenter and Edwards. Indeed, the evidence suggests that, at the temperature at which it is formed, it does not exist over a much larger area than that with which the present investigation is concerned.

Before passing on to the consideration of the alloys containing 2 and 3 per cent. of copper, reference may be made to those sections of the model which are parallel to the base, *i.e.*, those which represent the condition of the alloys at constant temperatures. These are shown in figs. 19–21. Fig. 19 represents a horizontal section or plan cut through the model at 400° C.

The next plan (fig. 20) represents the constitution of the alloys at 300° C., while fig. 21 shows the constitution at 200° C. It is interesting to note that there is no difference between the phases present in the field below the ternary eutectic and in that below the area containing $\beta-\epsilon$ only. This is due to the fact that there is β in both regions, which decomposes into $\eta+\gamma$, and there is also ϵ in both regions, which passes through the transformation temperature unaltered. While there is no phase difference in the alloys below about 285° C. in that part of the diagram where three phases are present, there is a difference in the arrangement of these phases, for, below the ternary eutectic, there are two types of η present, one which has separated from the liquid, and one which has arisen as a result of the decomposition of β, while, in those alloys which contain no ternary eutectic, only the second kind of η is present.

Finally, fig. 18 is a contour map of the liquidus surface. The valleys of the $\eta-\beta$ and the $\beta-\epsilon$ complexes, meeting at the ternary eutectic, can be readily distinguished.

The sections containing 2 per cent. of copper and 3 per cent. of copper have not yet been considered in this paper. In this connection, reference must be made to figs. 22 and 23 or to the model. In figs. 22 and 23, the line BC separates those alloys whose primary crystallisation is η from those which commence by depositing ϵ from the liquid. It will be seen that the lines representing alloys containing 2 and 3 per cent. of copper cut this line, so that those alloys in these series lying to the right of BC have ϵ as a primary separation, while, in those to the left, η is the first separation. In the former case, the separation of ϵ is followed by the reaction ϵ + liquid = η, but, if the composition of the alloy is represented by a point lying within the shaded area, BCJ, of fig. 23, there is not enough ϵ to complete this reaction, and the result is that there is only η + liquid left over. To the right of CJ we have $\eta + \epsilon$ + liquid.

In the case of the alloys with 2 per cent. of copper the reaction dies out very rapidly—probably on the addition of about 1 per cent. of aluminium, but in the alloys containing 3 per cent. of copper the reaction takes place over a much greater range of composition, ceasing at about 5 per cent. of aluminium. The evidence for the existence of this reaction is, perhaps, not very strong in the 2 per cent. copper series; all that can be said with certainty is that it exists in an alloy containing 98 per cent. of zinc and 2 per cent. of copper, but that the addition of quite a small amount of aluminium appears to eliminate it. In the case of the alloys containing 3 per cent. of copper, however, the evidence is much stronger. The reaction is indicated by the thermal curves of the alloys containing 2 per cent. of aluminium and probably also on those of 3Z4, and in alloys containing at least as much as 1·5 per cent. of aluminium, the microscope shows evidence of primary separation of ϵ.

In the alloys containing both 2 per cent. and 3 per cent. of copper there must be a very narrow area in which the $\beta-\eta$ complex solidifies, but there is this difference: that in the 2 per cent. series the lower boundary of the area represents the final solidification of this complex, while in the 3 per cent. series it is the line of solidification of the ternary eutectic. This will be seen more clearly when the sections representing alloys of constant aluminium-content are considered.

In the 3 per cent. copper series this area does not commence until about 5 per cent. aluminium is reached, as there is no separation of either primary η or β in the case of the alloys containing less than 5 per cent. of aluminium. The upper boundary of the area will meet the lower one at this composition, and also at a point (outside the diagram) which marks the limit of the separation of η. The lower limit of the area in the case of the 2 per cent. copper series is not well fixed, but it appears to be in the neighbourhood of 1 per cent. aluminium.

The lines separating the fields containing η+liquid from the fields containing $\eta+\epsilon$+liquid are shown dotted, as their exact position has not been determined.

The sectional diagrams representing alloys of constant aluminium-content are shown in figs. 10 to 17. Fig. 10, which represents the equilibrium conditions of the alloys of copper and zinc, has already been considered. The remaining sections can be divided into two groups: those containing up to 5 per cent of aluminium, where the primary separation at the zinc-rich end of the section consists of η, and those containing 7 per cent. of aluminium and over, where the primary separation consists of β. To make matters clearer, it will be well to consider a series of typical alloys in the sections containing, say, 2 per cent. of aluminium and 10 per cent. of aluminium. In the first of these sections (fig. 11) the alloy containing 1 per cent. of copper (1Z2) commences to solidify at about 407° C. by the separation of η, which continues until the temperature of about 380° C. is reached, when the binary complex of η and β commences to separate, this separation taking place over a range of about 2° C.

In the case of the alloy 6Z2 the first material to solidify is ϵ, the solidification commencing at 480° C. and continuing with falling temperature until the reaction ϵ+liquid $=\eta$ commences at 400° C. At 378° C. the ternary eutectic separates.

The depressed binary complex is shown on a larger scale in the inset diagram of fig. 11. It is not possible to separate these lines pyrometrically, but theoretically they must exist. The lowering of the ternary eutectic, compared with that of the binary, appears to be about 2° C and is just perceptible on most, if not all, the sections.

Alloys 2Z10, 3Z10, and 5Z10 may be taken as typical of the second group (fig. 15). In the case of 2Z10 and 3Z10 the first constituent to separate on cooling is β, the crystallisation commencing at 415° C. and 400 C. respectively, but in the case of 5Z10 solidification commences at 400° C. by the separation of ϵ. The β continues to separate in 2Z10 until the temperature of 379° C. is reached, when the $\beta-\eta$ complex is deposited over a very narrow range of temperature. When 5 per cent. of copper is present, β separates from 400° C. to 390° C. and the $\beta-\epsilon$ complex comes out of solution during the next 12° C., after which the remaining liquid solidifies as ternary eutectic. The alloy containing 8 per cent. of copper behaves in the same way (at somewhat different temperatures) except that the primary separation is no longer β but ϵ.

There remains the small group of alloys lying between the two types of section where the primary zinc-rich separations are β and η respectively. These are represented by the diagram showing the section containing 6 per cent. of aluminium (fig. 13). Here it will be seen that, in alloys containing less than about 2·5 per cent. of copper the primary separation consists of crystals of β; at about 2·5 per cent. of copper it consists of the pure $\beta-\eta$ complex; and between about 2·5 of copper and 3·5 per cent. of copper it consists of η. Higher copper-contents give rise to a primary separation of ϵ. The temperature ranges over which these reactions occur are in general too small to be shown by the thermal curves, but an examination of the constitution of the alloys containing 6 per cent. of aluminium on the various constant copper sections shows that the diagram must be similar to that given in fig. 13. This can also be seen from fig. 18, on which the various binary complex valleys are plotted. It will be seen that the line representing the alloys containing 6 per cent. of aluminium starts in the field where the primary separation is β, then crosses the $\beta-\eta$ complex valley, enters the primary η field, and finally passes into the primary ϵ area. This part of the diagram is plotted on a larger scale in the inset to fig. 13.

Reference has been made in several places to the reaction,

$$\beta = \eta+\gamma.$$

This reaction, which was first observed by Rosenhain and Archbutt in the aluminium-zinc series, occurs in all the alloys containing more than about 1 per cent. of aluminium, *i.e.*, in all alloys in which β exists, either as a primary separation or as one constituent of an eutectic or a complex. It was found that, even with the slow rates of heating and cooling adopted, there was a lag of about 20° C. to 25° C. between the temperature of the reaction as indicated by the heating curves and by the cooling curves. The points given by the heating curves are very much more regular than those given by cooling curves, and this, together with the fact that there is much less liability for an arrest point to show superheating on a heating curve than to show undercooling on a cooling curve, has decided the authors to regard the temperature shown on the heating curves as the correct one, rather than to attempt to take a mean.

Perspective drawings of the model are given in figs. 22, 23 and 24. In each of these, various planes are shaded so as to emphasise the different stages in the solidification of the alloys, and above each drawing is a plan of the shaded areas. In fig. 22 the liquidus surfaces are shaded. In any alloy whose composition is indicated by a vertical line passing through the area ABCD, solidification commences by the separation of η. Similarly, over the area CDEFG, the primary separation is β, while ϵ commences to separate in the area CGHB. The point C represent the constitution of the ternary eutectic, while CD and CG are the *loci* of the separation of the $\eta-\beta$ and the $\beta-\epsilon$ complexes, free from any other primary separation.

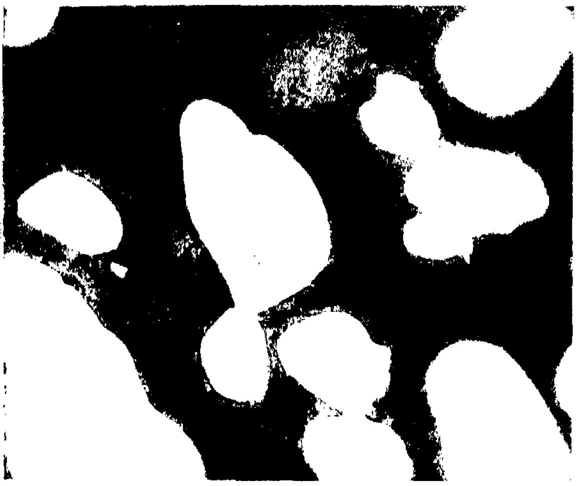
Fig. 23 represents the next stage in the solidification of the alloys. The $\eta-\beta$ complex commences to separate at a temperature indicated by the plane COPNM, the $\beta-\epsilon$ eutectic beginning to crystallise over the area CMLGK. The area CKIB marks the temperature at which η commences to solidify as a result of the reaction $\epsilon+\text{liquid}=\eta$. Fig. 24 represents the solidus surface of the system, which is divided into five areas which may be tabulated as below :—

No.	Area.	Description.
1	AJOP	End of solidification of η by itself.
2	JIQO	,, ,, in presence of ϵ.
3	OPNU	,, ,, of $\eta-\beta$ complex.
4	OUTRQ	The ternary eutectic plane.
5	RST	End of solidification of $\beta-\epsilon$ complex.

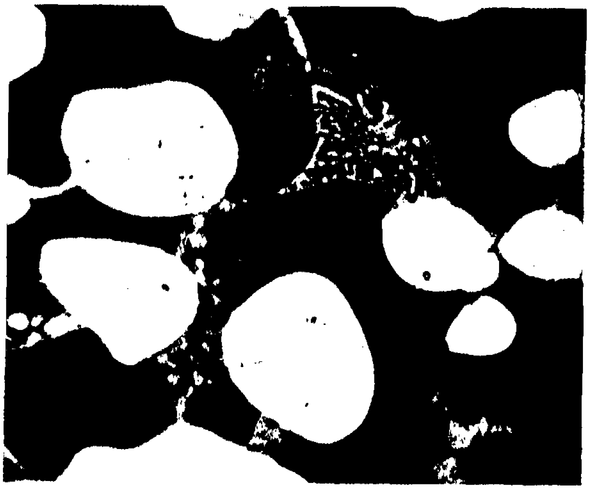
The area VWXY in fig. 22 indicates the temperature at which the decomposition of β takes place.

Afterword.

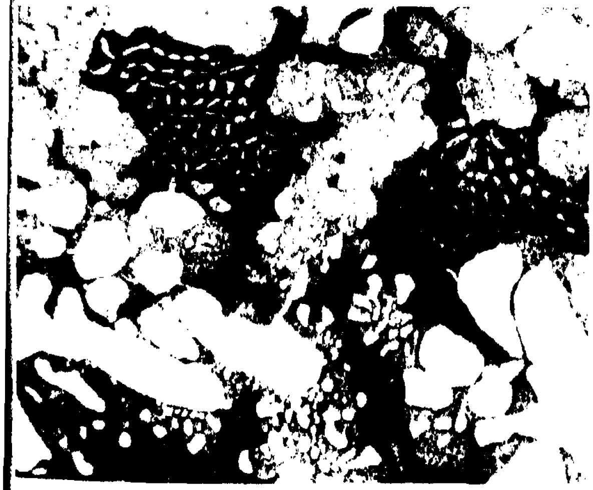
The authors wish to express how much they owe to Dr. W. Rosenhain, F.R.S., Superintendent of the Metallurgical Department of the Laboratory, for the interest he has shown in the work, for the manner in which he has always been willing to assist them with advice and suggestions, and especially for the advantages which they gained, both in the accuracy and in the expedition of the work, through using the apparatus and methods designed and installed in his Department of the National Physical Laboratory. To Mr. W. H. Withey, B.A., and Mr. P. G. Ward, B.Sc., A.R.C.S., of the Chemical Division, who carried out the majority of the analyses, and to others of their colleagues, they would express their indebtedness.



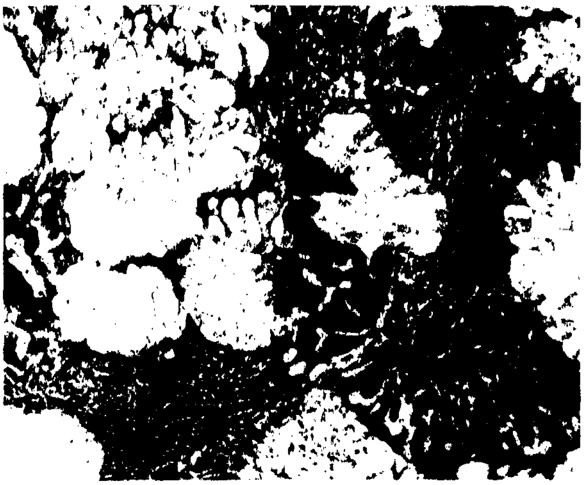
26



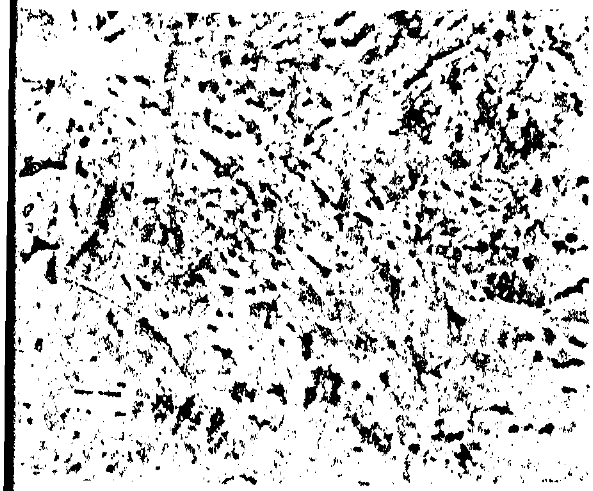
27



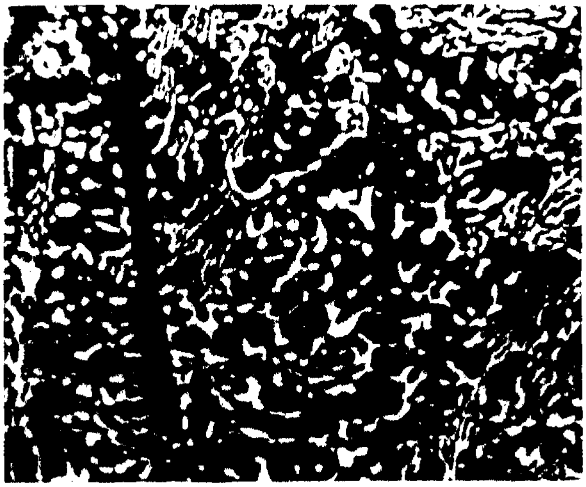
28



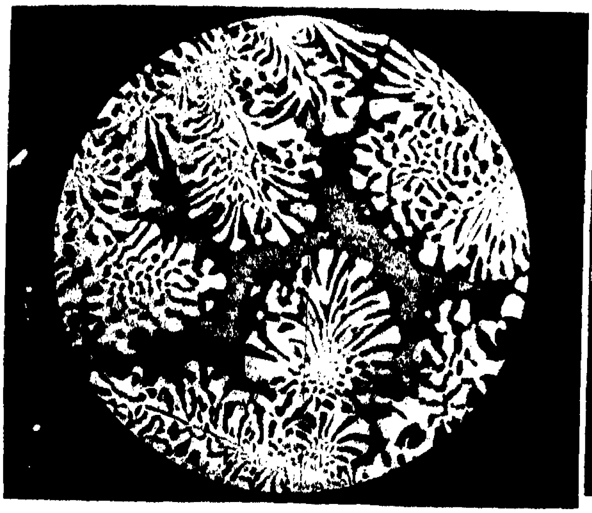
29



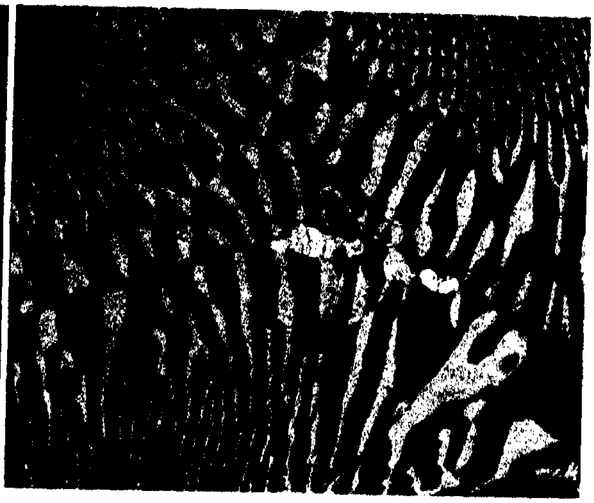
30



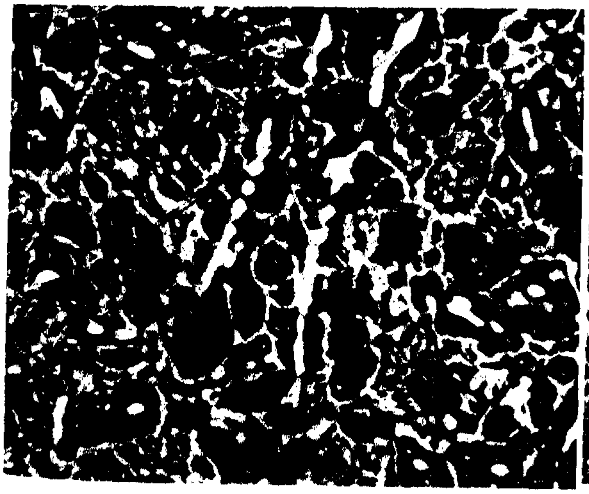
31



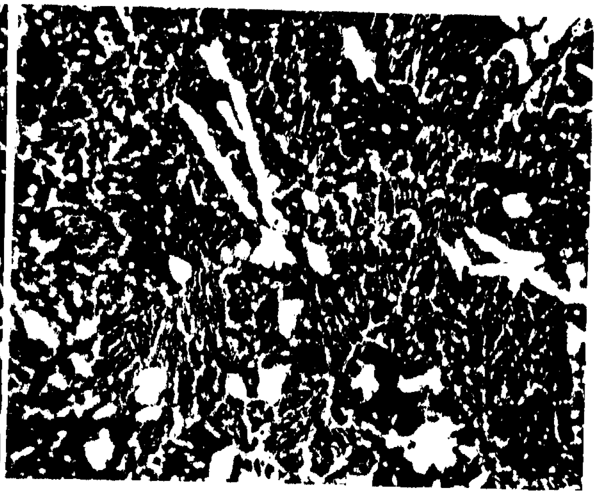
32



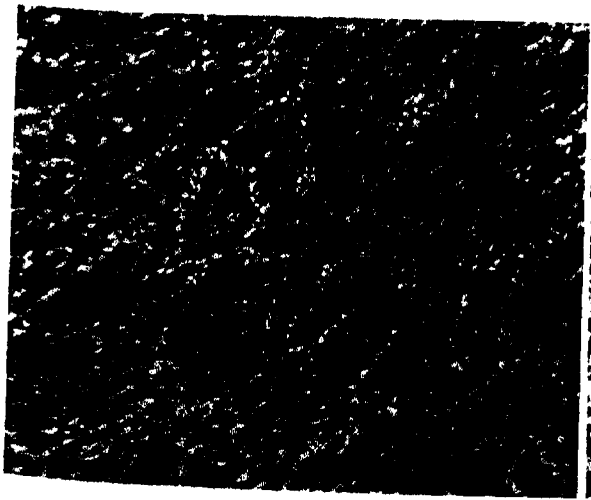
33



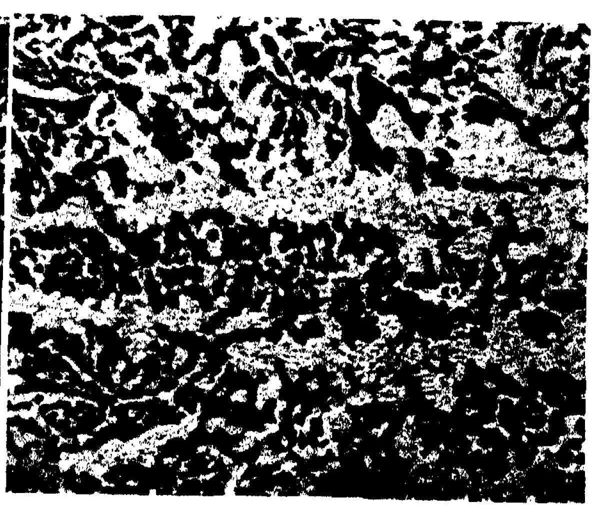
34



35



36



37

DESCRIPTION OF PLATES.

PLATE 1.

Fig. 26.—7Z½. Slowly cooled from above the liquidus to 345° C. × 300.
Fig. 27.—7Z2. Slowly cooled from above the liquidus to 345° C. × 300.
Fig. 28.—7Z8. Slowly cooled from above the liquidus to 345° C. × 300.
Fig. 29.—7Z10½. Slowly cooled from above the liquidus to 350° C. × 300.
Fig. 30.—7Z12. Annealed at 345° C. × 300.
Fig. 31.—7Z15. Annealed at 345° C. × 300.

PLATE 2.

Fig. 32.—9Z12. Slowly cooled from above the liquidus to 350° C. × 100.
Fig. 33.—3Z6. Slowly cooled from above the liquidus to 370° C. × 300.
Fig. 34.—4Z4. Annealed at 300° C. × 300.
Fig. 35.—4Z6. Annealed at 300° C. × 300.
Fig. 36.—4Z7. Annealed at 300° C. × 300.
Fig. 37.—4Z8. Annealed at 300° C. × 300.

The Vacuum Arc Spectra of Sodium and Potassium.

By SNEHAMOY DATTA, M.Sc. (Calcutta), D.I.C., Research Student, Imperial College of Science.

(Communicated by Prof. A. Fowler, F.R.S.—Received February 15, 1921.)

1. *Introductory.*

In connection with the investigation of spectral series, the need for more accurate measurements of the spectra of the alkali metals has long been felt, especially in the case of the diffuse series and the higher members of the sharp series. When it was found that the "sodium vapour lamp," recently described by Lord Rayleigh,* gave extremely narrow bright lines, such that the term diffuse as applied to one of the series becomes a misnomer, it seemed possible that the desired improvement in the measures of the wave-lengths might be attained by the use of this ideal vacuum arc as a source. Definite improvements in the measures for sodium, as proved by the constancy of the separations of pairs of lines, suggested that a potassium vapour lamp might also be used with probable success. The potassium lamp was much more troublesome than the sodium one, and the lines of the diffuse series were less sharply defined. The new measures of potassium lines, however, are probably more exact than before, and an interesting combination pair, indicating satellites to the diffuse series has been observed.

---

* R. J. Strutt, Bakerian Lecture, 'Roy. Soc. Proc.,' A, vol. 96, p. 272 (1919).

2. *Experiments.*

The sodium vapour lamp was excited precisely in the manner described by Lord Rayleigh, being run on a 200-volt circuit, with a current of about 2·5 ampères. When once started it would continue even when the process of evacuation by the Gaede pump was stopped, and heating the bottom bulb by the ring burner done away with, the vapour then being produced by the heat of the arc only; but it was noticed that an increase of vapour pressure caused the lines to become slightly diffuse, and the pump was therefore run continuously during the exposures of the plates.

With potassium, however, continuous evacuation caused considerable difficulty. The vapours rising up by the upward suction of the pump condensed on the sides of the upper bulb and trickled down the stem, rendering it either partially opaque or making a short circuit along the side by the deposited metal. After various trials it was found best to raise the temperature of both the bulbs and stem in such a way that the deposition of metallic potassium was impossible. For this purpose the upper bulb was heated by a ring burner, having a tiny flame, whereas the bottom one was heated by a Bunsen burner intermittently. Even then it was found desirable to decrease the vapour pressure inside the bulb in such a way that the temperature equilibrium was not seriously disturbed by a sudden uprush of the vapour caused by the pump. A side reservoir was therefore connected to the ordinary pump circuit by means of a T-piece. After evacuating the bulb with the pump, the pump circuit was closed, and the side reservoir was kept immersed in liquid air; in this way the pressure could be kept down without making the lamp unstable. The potassium lamp thus required continual attention, but otherwise it was quite satisfactory and gave a sufficiently intense light.

As the bulb was heated it emitted a white light, and the subordinate series lines, as well as the first pair of the principal series lines, appeared to become stronger at the expense of the rest of the members of the principal series, and exactly the opposite took place when it was cooling, the bulb then appearing distinctly bluish. The combination line $1s-2d'$ is fairly persistent all the time.

Preliminary photographs were taken with a spectrograph of moderate dispersion, and final measurements were made on photographs taken in the first order, and in some cases in the second order spectrum of the

10-foot concave grating at the Imperial College, giving a dispersion of about 5·6 A.U. per millimetre in the first order.

To get the higher members of the subordinate series it was necessary to give an exposure of about four hours, and in order to guard against possible shifts due to change of temperature or to mechanical displacements during the exposure, the iron comparison spectrum was taken both before and after exposure to the lamp. The appearance of the iron lines in the photograph showed that there was no such shift. The standards of reference were iron lines selected from the table given by Burns.

3. *The Line Spectrum of Sodium.*

The new observations of sodium are compared with those of Kayser[*] and Zickendraht[†] in Table I, which includes the separation $\Delta\nu$ of the doublets (in wave numbers), and shows the series to which each of the pairs belongs.

The lines of the sharp series are closely represented by a formula of the Hicks' type, viz.:—

$$\left.\begin{array}{l} S_1(m) = 24476{\cdot}18 \\ S_2(m) = 24493{\cdot}35 \end{array}\right\} - \frac{109678{\cdot}3}{(m + 0{\cdot}650835 - 0{\cdot}013290/m)^2},$$

the constants of which were calculated from the less refrangible components of the pairs 3, 4, and 6. The difference between the observed and calculated wave numbers are also included in the table. The variable parts or "terms" of the other series are then as indicated in the table, their limits being obtained from certain well-established laws of series spectra, according to which the limits of the diffuse series may be taken to be identical with those of the sharp, and the limit of the principal series may be obtained by adding the wave numbers of the first principal lines to the respective limits of the diffuse and sharp series.

4. *The Band Spectrum of Sodium.*

An interesting feature of the sodium vapour lamp was the production of a channelled space emission spectrum, which had its origin only in the bulb of the lamp. These bands always made their appearance when the arc was first started and persisted so long as the bulb did not attain too high a temperature. On removing the ring burner so as to decrease the temperature of the bulb, some of the previous heads only appeared without the entire flutings. A photograph with an instrument of sufficient dispersion to be of value for measurement would have been difficult to obtain on account of

---

[*] Kayser, 'Handbuch der Spectroscopie,' Band VI, p. 116.

[†] H. Zickendraht, 'Ann. d. Phys.,' vol. 31, p. 233 (1910).

Table I.—Series Lines of Sodium.

Kayser and Runge. λ (Row.)	Δν.	Datta. λ : I.A.	ν (vac.).	Δν.	m.	mp.	$\nu_o - \nu_c$.
		Principal series :	$1s - mp_1$ $1s - mp_2$	$1s = 41449 \cdot 53$			
5890·19 96·16	17·18	5889·963 95·982	16973·35 ·56·17	17·18	(1)	24476·18 93·36	
		Sharp series :	$1p_1 - ms$ $1p_2 - ms$	$1p_1 = 24476 \cdot 18$ $1p_2 = 24493 \cdot 36$		*ms* (mean).	
6161·15 54·62	17·22	6160·725 54·214	16227·37 44·54	17·17	(3)	8248·81	*
5153·72 49·19	17·07	5153·645 49·090	19398·34 415·51	17·17	(4)	5077·84	*
4752·19 48·36	16·97	4751·891 48·016	21038·37 55·55	17·18	(5)	3437·81	+0·17
4546·03 42·75	15·89	4545·218 41·671	21995·00 2012·18	17·18	(6)	2481·18	*
Zickendraht. λ : I.A.							
4423·2 19·9	16·88	4423·31 19·94	22601·16 18·40	17·24	(7)	1875·00	−0·38
4344·7 41·4	17·49	4344·8 41·5	23009·5 27·1	17·5	(8)	1466·47	−0·50
		Diffuse series :	$1p_1 - md$ $1p_2 - md$	$1p_1 = 24476 \cdot 18$ $1p_2 = 24493 \cdot 36$			
K. & R. (Row.).						*md* (mean).	
5688·26 82·90	16·59	5688·222 82·675	17575·30 92·47	17·17	(3)	6900·88	
4983·53 79·30	17·04	4982·864 78·608	20063·20 80·34	17·14	(4)	4413·00	
4669·4 65·2	19·26	4668·597 64·859	21413·78 30·89	17·16	(5)	3062·46	
4500·0 494·3	18·19	4497·724 94·266	22227·11 44·25	17·14	(6)	2249·09	
Zickendraht. λ : I.A.							
4393·5 90·1	17·62	4393·45 90·14	22754·77 71·96	17·19	(7)	1721·41	

* Used in calculation of constants.

the low intensity of the bands, and only rough measurements were therefore made with a constant-deviation spectrometer. The heads of the bands appeared in pairs and the bands themselves faded off towards the violet. These bands show very little agreement with the previously recorded absorption or fluorescent bands.* They can, however, be divided into several groups having a nearly constant frequency difference between the heads, thus:—

---

* Roscoe and Schuster, 'Roy. Soc. Proc.,' vol. 22, pp. 362–364 (1864); Wood and Moore, 'Astro. Phy. J.,' vol. 18, pp. 94–111 (1903); Wood and Galt, 'Astro. Phy. J.,' vol. 33, pp. 72–80 (1911).

Group I	λ6688 – 6246	7 equally spaced heads of bands.		
Group II	λ6182 – 5756	9	„	„
Group III	λ5554 – 5340	7	„	„
Group IV	λ5281 – 5101	7	„	„

5. *The Line Spectrum of Potassium.*

As in the case of sodium, the new observations of potassium are compared with those of Kayser and Runge[*] in Table II, which includes the separation $\Delta\nu$ of the doublets (in wave numbers).

The lines of the sharp series may be closely represented by a formula of the Hicks' type, viz.:

$$\left.\begin{array}{l} S_1(m) = 21963{\cdot}06 \\ S_2(m) = 22020{\cdot}84 \end{array}\right\} - \frac{109678{\cdot}3}{(m+0{\cdot}821947-0{\cdot}054486/m)^2},$$

for which the less refrangible components of the pairs 5, 6, and 7 were used in the calculation of constants. The residuals are shown in Table II. From the limits $1p_1$, $1p_2$, thus determined, the limit of the principal series was also derived in the usual way, and the variable parts or "terms" *mp*, *ms*, and *md* of all the lines were then calculated as shown in the table.

6. *The Second Principal Pair of Lines.*

In connection with combination terms involving the second principal pair of potassium, Paschen found it necessary to adopt wave-lengths some tenths of an Ångström higher than those measured by Kayser and Runge, and it was therefore thought desirable to investigate this pair more completely.

Several photographs were taken of this pair as given by the arc in air as well as in the vacuum bulb. The measures obtained from different photographs of the bulb were in close agreement with each other. But those obtained from the arc in air differed according as the lines were reversed or unreversed. In ordinary circumstances the lines of the principal series are strongly reversed, but by feeding the arc sparingly with potassium bicarbonate it was found possible to obtain plates with the lines unreversed. Measurements showed that when the reversal was broad the wave-length of the centre of reversal was about 0·1 A.U. higher than that of the unreversed line, indicating that reversal takes place unsymmetrically.

A further interesting point observed was that when the reversal is narrow the centre has a distinctly lower value than when it is broad, suggesting that

---

[*] H. Kayser, 'Handbuch der Spectroscopie,' Band V, p. 634.

Table II.—Series Lines of Potassium.

Principal series: $1s - mp_1$, $1s - mp_2$ $1s = 35005·88$

Kayser and Runge. λ (Row.).	Δν.	Datta. λ : I.A.	ν (vac.).	Δν.	m.	$mp_{1,2}$.	$\nu_o - \nu_c$.
—	—	7664·94* 99·01*	13042·82 12985·04	57·78	(1)	21963·06 22020·84	
4044·29 47·36	18·76	4044·140 47·201	24720·18 24701·49	18·69	(2)	10285·70 10304·39	
3446·49 47·49	8·41	3446·722 47·701	29004·70 28996·55	8·15	(3)	6001·18 6009·33	
3217·27 17·76	4·73	3217·01 17·50	31075·88 71·05	4·83	(4)	3930·00 3934·83	
		Sharp series: $1p_1 - ms$ $1p_2 - ms$		$1p_1 = 21963·06$ $1p_2 = 22020·84$		*ms* (mean).	
6938·8 11·2	57·56	6939·50 11·80	14406·35 64·10	57·80	(3)	7556·71	+0·50
5802·01 782·67	57·04	5802·16 782·77	17230·21 87·97	57·77	(4)	4782·85	+0·11
5340·08 23·55	58·14	5339·670 23·228	18722·56 80·40	57·84	(5)	3240·50	†
5099·64 84·49	58·42	5099·180 84·212	19605·54 63·28	57·74	(6)	2357·52	†
4956·8 43·1	55·91	4956·043 41·964	20171·61 229·34	57·63	(7)	1791·45	†
—	—	4863·61 49·88	20555·16 613·33	58·17	(8)	1407·90	+0·29
—	—	4800·16 786·89	20826·82 984·58	57·76	(9)	1136·24	−0·55
—	—	4754·5 41·6	21026·7 21084·0	57·3	(10)	936·36	−1·07
		Diffuse series: $1p_1 - md$ $1p_2 - md$		$1p_1 = 21963·06$ $1p_2 = 22020·84$		*md* (mean).	
5832·23 12·54	58·08	5832·31 12·71	17141·13 98·91	57·78	(4)	4821·93	
5359·88 43·35	57·72	5359·521 42·974	18653·20 711·00	57·80	(5)	3309·86	
5112·68 097·75	57·28	5112·204 097·144	19555·61 613·38	57·77	(6)	2407·45	
4965·5 52·2	54·09	4965·038 50·816	20135·23 93·08	57·85	(7)	1827·88	
—	—	4869·70 56·03	20529·46 87·23	57·77	(8)	1433·60	
—	—	4805·19 791·08	20805·03 66·33	61·30	(9)	1158·03	
—	—	4759·31 45·58	21005·58 66·38	60·80	(10)	957·48	

Combination.			ν (cal.).	Series.
4642·172 41·585	21535·62 38·36	2·74	21535·46	$1s - 2d$ $1s - 2d'$

* These values have been taken from Meggers.

† Used in calculation of constants.

the reversed line also shrinks unsymmetrically. The same thing is suggested by the sharp and diffuse series lines emitted by the vacuum bulb, for they shrink into fine lines which have in some cases several tenths of an Ångström unit lower value than those given by Kayser for the arc in air.

The following table shows the various measures of the pair :—

Table III.

Source.	λ I.A. p_1 (2).	λ I.A. p_2 (2).
Vacuum arc (2nd order) …………	4044·140	4047·201
„ „ (1st „) …………	4044·146	4047·206
Arc in air, unreversed ……………	4044·160	4047·230
„ fine reversal …………	4044·168	4047·239
„ broad reversal ………	4044·274	4047·342

7. *The Combination Pair* 1*s*–2*d*, 1*s*–2*d*′.

The close pair of narrow bright lines at 4642·172, 4641·585 has been previously observed only as a single line but has been recognised as the combination 1*s*–2*d*. The occurrence of a *pair* of lines at once suggests the combinations 1*s*–2*d* and 1*s*–2*d*′; or in other words that there is a satellite to the line d_1 (2) of the diffuse series, with a separation of 2·74 in wave number. This *d* (2) pair is in the infra-red, at wave-lengths 11771·73, 11689·76 according to Paschen, and the satellite has not yet been observed. The fact that the less refrangible component of the combination pair is the stronger would seem to indicate that the satellite to d_1 (2) is abnormally placed on the more refrangible side of the chief line. The probable existence of a satellite to d_1 (2) further suggests that the associated fundamental series, which has its lines in the far red and infra-red, should also consist of pairs, with separation 2·74, but these lines have not yet been recorded as doublets.

It does not seem possible to trace any definite relation between the satellite separation of potassium and those of rubidium and cæsium. The existence of satellites in the diffuse series of cæsium is well established, and their separations in successive terms of the series appear to be quite normal. In the case of rubidium, however, though an apparent satellite with a separation of 2·93 has been observed in the second member of the series *d* (3), there is no record of a satellite to the infra-red member *d* (2). Prof. Hicks* doubts the reality of the satellite to *d* (3), and the absence of a satellite to *d* (2) is rendered probable by the fact that the combination 1*s*–2*d* of rubidium (λ5165·18) only appears as a single line.

---

* Hicks, 'Phil. Trans.,' A, vol. 210, pp. 57–111 (1910).

8. *Sodium and Potassium Lines in the Solar Spectrum.*

Sodium is represented in the Sun by several coincidences, among which the most important lines are the lines D_1 and D_2 which constitute the first pair of the principal series. A recent publication of the Allegheny Observatory* shows that the lines with wave-lengths 8183·30 A. and 8194·82 A. which form the first pair of the diffuse series, are prominent in the infra-red solar spectrum. The new measures of sodium suggest additional identification of solar lines as shown in the following table:—

Table IV (*a*).—Sodium Lines in the Sun.

Series.	Rowland (λ).	Intensity.	Datta (λ.I.A.).	Correction to International scale.	Remarks.
s_1 (3)	6160·956	3	6160·725	+0·22	Previously recorded, but the new measures give much better agreement.
s_2 (3)	54·438	2	54·214		
d_1 (3)	5688·436	6	5688·222	+0·22	
d_2 (3)	82·869	5	82·675		
s_1 (4)	5153·848	000	5153·645	+0·18	C? (Rowland).
s_2 (4)	49·267	000	49·090		
d_1 (4)	4982·994	2	4982·864	+0·18	Masked by Fe.
d_2 (4)	78·785	3	78·608		
d_1 (5)	4668·749	1	4668·597	+0·17	Masked by Cr.
d_2 (5)	—	—	64·858		
d_1 (6)	4497·842	0	4497·724	+0·17	
d_2 (6)	94·356	00	94·266		

The existence of potassium in the Sun was considered uncertain until recently, because its presence was indicated by only one faint line, 4044·140 A. [$p_1(2)$]. The presence of the first pair of principal lines 7664·94 A. and 7699·01 A. was established in the course of the present investigation, and confirmatory evidence has been published by Meggers.†

The less refrangible member of the first pair, 7699·01 A., is unmistakably in the Sun and shows a shift at the Sun's limb of about 0·06 A. as compared with the centre, on a solar rotation plate kindly lent by Prof. Sampson, the calculated shift being about 0·058 A. The second line of the pair, 7664·94 A., is masked by a strong broad line in the A band of oxygen, and, although the rotation shift did not separate it in the photograph, it was found to produce a distinct broadening of the oxygen line by about 0·04 A. Lester's‡ value of this oxygen line is 7665·14 A., and, as it has a breadth of about 0·26 A., the potassium line is to be expected at

---

* Meggers, 'Publications of the Allegheny Observatory,' vol. 6, No. 3.

† Meggers, *loc. cit.*

‡ Lester, 'Astro. Phy. J.,' vol. 20, p. 81.

7665·14 + 0·13 − 7665·25 = 0·02 A., inward from the edge. Consequently, a shift of 0·06 A. will clearly broaden the line by 0·04 A. Meggers has obtained quite definite evidence on this point, by observing the notched appearance of the oxygen line in the juxtaposed spectra of the two limbs.

Other lines of potassium possibly represented in the solar spectrum are shown in the following table :—

Table IV (*b*).—Potassium Lines in the Sun.

Series.	Rowland (λ).	Intensity.	Datta (λ.I.A.).	Correction to International scale.	Remarks.
p_2 (1)	7699·297	—	7699·01	+0·28	Values taken from Meggers.
p_1 (1)	65·248	—	64·94		
d_1 (8)	6965·320	000	6965·0	+0·25	Values taken from Saunders.
d_2 (3)	36·290	0000	36·0		
d_1 (4)	5832·490	000	5832·31	+0·22	
d_2 (4)	12·942	0000	12·71		
d_1 (5)	5359·714	000	5359·521	+0·19	
d_2 (5)	43·148	0000	42·974		
p_2 (2)	4047·338	00	4047·201	+0·15	Previously recorded.
p_1 (2)	44·294	0	44·140		
p_1 (3)	46·857	000	3446·722	+0·14	

9. *Summary.*

1. The wave-lengths of the sharp and diffuse series of lines of sodium and potassium have been carefully measured in a source giving well-defined lines.

2. Channelled-space emission bands have been observed in the sodium vapour lamp, and their heads approximately measured.

3. The combination line $1s-2d$ in potassium has been resolved into a pair, and their wave-lengths measured.

4. The presence of potassium in the Sun has been established, and some additional sodium lines have been identified with solar lines.

In conclusion, I beg to thank Prof. A. Fowler, F.R.S., for his taking helpful interest in the work, Lord Rayleigh for the use of his lamp, and Prof. Sampson for kindly lending me a Solar rotation plate.

On Spectrophotometry in the Visible and Ultra-violet Spectrum.

By T. R. MERTON, M.A., D.Sc., F.R.S., Professor of Spectroscopy in the University of Oxford.

(Received February 15, 1921.)

In a previous communication* the use of the neutral wedge for the determination both of the photographic and absolute intensities of spectrum lines has been discussed, and the results of its application to the study of certain phenomena relating to the spectra of hydrogen and helium have been given.† The method has been found to be simple and convenient for the study of the relative intensities and of the structure of broadened lines in the visible spectrum. There are, however, certain features of the method which under some circumstances limit its application; in particular the density of the "neutral" wedge increases with the wave-number, and for the investigation of the ultra-violet down to $\lambda = 2000$A, the method is at present inapplicable. Although it might be possible to construct wedges of crown glass with a very small angle for use in the ultra-violet, the fact that a number of such wedges would be required for different ranges would destroy the principal advantage of the method. Whilst no special precautions are required in the investigation of extended sources of light, it is evident that when small sources of light are used it is of the utmost importance to ensure that that portion of the slit which is behind the wedge is uniformly illuminated before the wedge is put into position, and in the case of selected regions of a source of light, *e.g.*, the spectrum from a particular point in the electric arc, it is only possible to obtain correct results by means of devices which entail a very considerable loss of light.

The present communication relates to a method which, whilst not superior to the wedge method in accuracy, has the advantage that it can be used in any part of the spectrum which can be photographed through quartz lenses and prisms, and its application to the extreme ultra-violet beyond $\lambda = 2000$A should present no serious difficulty. The method consists in crossing the dispersing system, *e.g.*, the prism of the spectrograph, with a very coarse grating, and reducing the length of the slit to a very small value. The grating is inserted between the prism and the camera lens of the spectrograph with the rulings perpendicular to the refracting edge of the prism, and a continuous spectrum thus appears on a plate as a dark central strip with a succession of other strips of different intensities on either side, the intensities of these orders being determined by the ruling of the grating and the width of the strips by the length of the slit. In the case of a discontinuous spectrum the "lines" are found to consist of dots of different intensities on either side of the central dot. It is evident that if the last dots which are just visible in the case of two lines are noted, a previous knowledge of the relative intensities of the different orders corresponding to these dots at once enables the relative intensities of the lines to be determined. It will be noted that the fact that the slit is reduced to a point not only enables different regions of a light source to be investigated without difficulty, but suggests a further application, which will be considered later, in the study of stellar spectra.

---

* Merton and Nicholson, 'Phil. Trans.,' A, vol. 217, p. 237 (1917).

† *Cf.* 'Phil. Trans.,' A, vol. 220, p. 137 (1919); 'Roy. Soc. Proc.,' A, vol. 98, p. 255 (1920).

The experiments recorded in this communication have been made with a Hilger quartz spectrograph, with which the spectrum from the extreme red down to $\lambda = 2000$ A. is recorded on a plate 10 inches in length, and also with a constant deviation spectroscope with a camera attachment. Lord Rayleigh* has shown that in the case of a grating consisting of transparent bars of width, a, alternating with opaque bars of width b, the brightness of the mth order B_m is given by the expression

$$B_m : B = [1/m^2\pi^2] \sin^2 [am\pi/a+b],$$

where B would be the brightness of the central image if the whole of the grating were transparent; but it is probably impossible to produce gratings which conform absolutely to this ideal condition, and in my experiments all the gratings which were tried, whilst exhibiting qualitatively the distribution of intensity to be expected according to theory, showed anomalies of various kinds, which were due no doubt to irregularies in the ruling.

It is necessary, therefore, to calibrate the gratings, that is to say, to determine experimentally the distribution of intensity in the different orders. A number of gratings have been made and tested, and the most convenient ruling for use with the quartz spectrograph has been found to contain about twenty-five lines to the inch. Wire-wound gratings are useless for the present purpose, for a very small rotation of the grating about an axis parallel to the rulings completely alters the distribution of intensity in the different orders, by changing the ratio of the transparent to the opaque parts of the grating. It is also essential that the gratings should not be made by any method which would involve cutting a groove in a transparent surface, for in this case we should have, in addition to the

---

* 'Collected Papers,' vol. 1, p. 213.

above-mentioned anomalies, which are due to irregularities in the ruling, the still more disturbing effects, which are not independent of the wave-length, due to the form of the groove.

The method adopted* has been to coat a quartz plate with a very thin layer of lamp-black by holding it over a dish of burning toluene, and to flow over the surface when cold a small quantity of alcohol containing a very small trace of shellac, which seems to improve the sharpness of the edges when the grating is ruled. The ruling was performed with a bone tool on a shaping machine provided with an automatic feed, and the resulting grating was experimentally calibrated.

The calibration was carried out with the aid of a neutral wedge, for which the density step, as a function of the wave-length, had been determined according to the method described in the previous investigation;† an essentially similar method for determining the density of a medium in the form of a wedge has since been described by Slade and Toy.‡ For this purpose the constant deviation spectrograph was used, and the grating was set with the rulings parallel to the refracting edge of the prism, with the wedge set in the usual manner in front of the slit, which for this purpose was of such a length as to cover the whole length of the wedge. As a source of light, a mercury arc lamp was used, and individual radiations, *e.g.*, $\lambda = 4359$ A., were isolated by inserting absorbing screens in front of the lamp and by an appropriate choice of photographic plates.

The resulting pattern on the plate is shown in fig. 2*a*, in which the second groups of orders, beyond the points at which the values of $[am\pi/a+b]$, in the expression given above, become greater than π, are clearly visible. The relative intensities of the different orders were determined by measuring the heights of the orders on the plate, the relative intensities being proportional to $\log 10^{-1}[d_\lambda h_\lambda/m]$, where d_λ is the change of density per millimetre on the wedge, h_λ the height in millimetres of an order on the enlargement, and m the magnification of the slit on the enlargement. If the intensity of the central order be taken as unity, and if the intensity of the mth order be M, it is evident that, if a spectrum line, λ_1, gives an image on which the central order is just visible, and the image of another line, λ_2, is just visible in the mth order, the ratio of their intensities, I_{λ_2} and I_{λ_1}, will be given by $I_{\lambda_2}/I_{\lambda_1} = 1/M$. A calibration curve is given in fig. 1, in which values of $\log I$ are plotted against the corresponding orders. The arrow denotes the central image, the intensity of which has been made equal to unity, and the ordinates thus give the logarithms of the intensities of lines which are just visible in corresponding orders on the abscissa scale.

---

* *Cf.* Wood, 'Physical Optics,' p. 211.

† *Loc. cit.*, p. 241.

‡ 'Roy. Soc. Proc.,' A, vol. 97, p. 191 (1920).

It is evident that the method is subject to the inconvenience that the scale of intensities is discontinuous, and in this respect it is inferior to the wedge method in its application to discontinuous spectra, though it is believed that estimates of the visibility of the last dot can attain a considerable degree of precision, since all such estimates are carried out over the same range of photographic density, and the asymmetry in the intensities of the orders on opposite sides of the central image, which was found on all the gratings, and which is clearly shown in fig. 1, is, in fact, an advantage, owing to the increase in the number of points available. It is also evident that, by a succession of exposures of different duration, the relative intensities of different lines can be linked up, in cases in which a single exposure fails to give sufficiently complete information. In its application to discontinuous spectra, the advantage of the method lies in the fact that it can be used in the ultra-violet, and photographs have been obtained over the whole range of wave-lengths, which could be photographed with the quartz spectrograph.

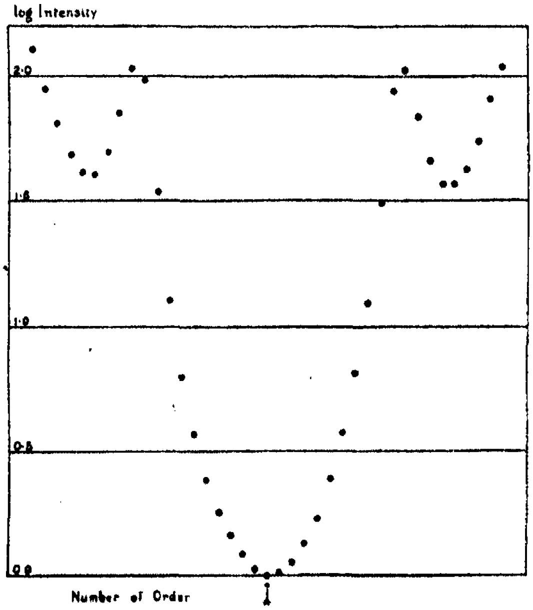


FIG. 1.

In the previous investigation* it has been shown that photographic intensities can be reduced to absolute values by adopting the positive crater of the carbon arc as a standard source of radiation, having a known distribution of intensity, the variation in the sensibility of the photographic plate for different wave-lengths and absorption of light by the spectrograph being thus eliminated. In fig. 2 is shown the pattern given by the positive crater of the carbon arc on a panchromatic plate, using the glass spectrograph with (B) the neutral wedge, and with (C) the grating, the helium lines being superposed as reference lines. In the visible portion of the spectrum the relation between the sensibility of panchromatic plates and the wave-length is of so complicated a character that a considerable number of exposures by the grating method might be required to elucidate it fully, and in this region it is more convenient to determine this relation (in which absorption by the spectrograph is included) by the use of the wedge; in the ultra-violet, where the relation of sensibility to wave-length is less complex, one or two exposures are sufficient to furnish all the information required. This relation must, of course, be determined for the particular plates which are used; with regard to the correction for dispersion when continuous and discontinuous spectra are compared, the reduction of photographic to absolute intensities, and other relevant considerations, reference may be made to the previous communication (*loc. cit.*).

---

* *Loc. cit.*, p. 242.

The grating method appears to be more particularly applicable to the study of the distribution of intensity in continuous spectra, the structure of broadened lines, and the quantitative investigation of absorption spectra, for in these measurements the discontinuity of the calibrated intensities is of no consequence. The measurement here consists in observing at what wave-lengths the strip due to each order is just visible, and there are thus a considerable number of points each of a specified intensity which can be plotted against the wave-length. This is shown in fig. 2, C; the appearance of an absorption band (on one side of the minimum) is precisely similar to the pattern at either end of the spectrum in the figure, the violet end of the spectrum being in this particular case determined by the absorption of the glass prism, and the red end by the sensibility of the photographic plate. If a source of light of constant intensity, giving a continuous spectrum, is used, it is sufficient to take two exposures of the same length on adjacent portions of the same plate, interposing in the case of one of these exposures a layer of an absorbing substance under investigation in front of the slit; these two exposures yield the information required for determining the density of the medium as a function of the wave-length.

In a recent investigation, H. H. Plaskett* has applied the wedge method to the study of the continuous spectrum and of the bright hydrogen lines in the star γ Cassiopeiæ, using a single prism spectograph attached to the 72-inch reflector of the Dominion Observatory, and has obtained interesting results relating to the black body temperature of the star. The wedge was placed in front of the slit, as in laboratory experiments, and uniform illumination was secured by allowing the image of the star to trail forward and backward over the slit.

---

* 'Monthly Notices Roy. Astr. Soc.,' vol. 80, No. 9, p. 771 (1920).

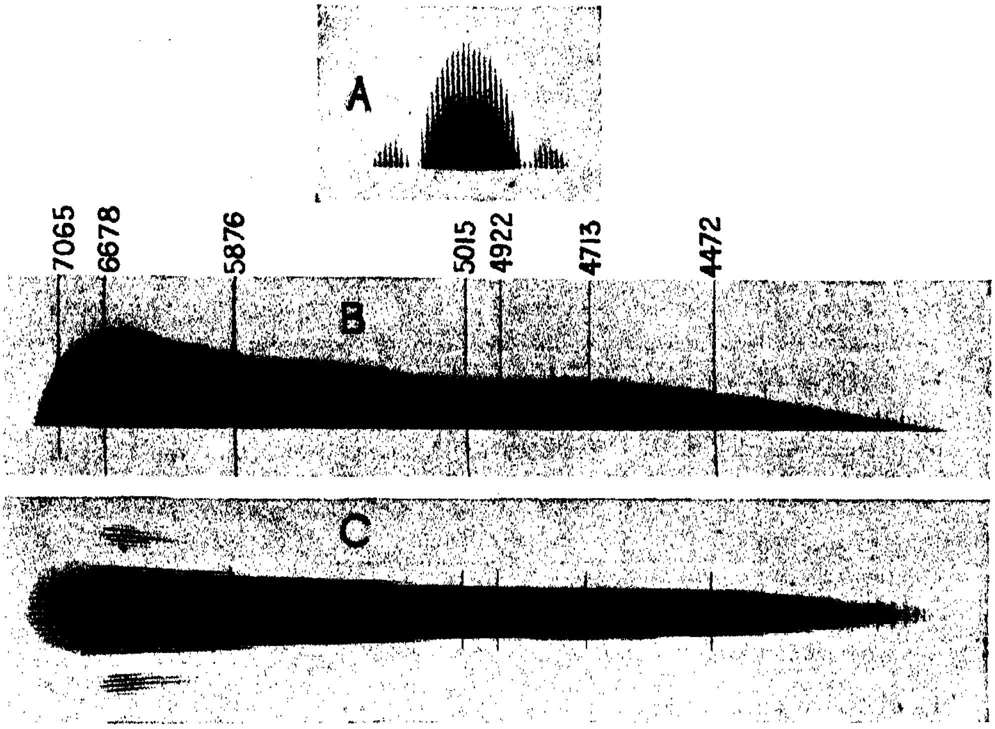


FIG. 2.

It would appear that the wedge method can only be applied to the study of stellar spectra by a procedure essentially similar to that adopted by Plaskett, and that instruments of very great light-gathering power are necessary; in Plaskett's experiments exposures up to two hours were given for γ Cassiopeiæ which is a star of photographic magnitude 2·01. I would suggest that the grating method is particularly applicable to the investigation of stellar spectra and that valuable results relating to the black body temperatures of stars might be obtained with telescopes of very moderate dimensions with the aid of an objective prism. In this case the star itself provides the beam of parallel rays which are produced in the laboratory by means of the collimator with the slit reduced to a pin-hole, and the objective prism is therefore directly applicable. It is only necessary to interpose a grating of suitable ruling between the objective prism and the objective in order to obtain a spectrum photograph of a star essentially similar to that shown in fig. 2, C. It would appear that the exposures required for this purpose would not be unduly long.

The loss of light with an objective prism is very much less than when a slit spectroscope is used, and the common procedure of allowing the star to trail slowly in a direction perpendicular to the dispersion in order to widen the spectrum, is for the present purpose to be expressly avoided, with a consequent gain in brightness, the spectrum being in this case widened by the action of the grating. It is also clear that the exposure need only be of such a length as to register the faintest portions of the spectrum on the central image, the brighter portions being thus visible in the higher orders. From the expression $B_m : B = [1/m^2\pi^2] \sin^2 [am\pi/a+b]$ it follows that the brightness of the central order $B_0 : B$ will be equal to $[a/a+b]^2$, and in the case of a equal to b the exposure would only be four times as long as it would be without the grating and perhaps considerably shorter than it would be if the image of the star were allowed to trail to any considerable extent; but a smaller value of a/b would probably be necessary for the study of a number of phenomena such as the structure of the nova bands and the time of exposure would accordingly be greater.

A Superior Limit to the Age of the Earth's Crust.

By HENRY NORRIS RUSSELL, Professor of Astronomy in Princeton University.

(Communicated by Mr. J. H. Jeans, Sec. R.S. Received February 26, 1921.)

Considerations based upon radio-activity have often been employed to obtain estimates of the ages of minerals, and of the duration of geological time; but it does not appear to have been remarked that similar arguments can be used to assign a maximum age for the existing crust of the Earth.

Uranium is continually disintegrating with a "half period" of 5×10^9 years, and no process is known by which the supply can be renewed. It is probable that the proportion of this element in the Earth's crust was greater formerly than at present—the computed amount necessary to account for the survival of that which now exists increasing exponentially with the time interval, a thousand fold in ten half periods and a million fold in twenty.

The existing proportion of uranium may be derived from that of radium, which, according to Joly,[*] averages about $2{\cdot}5 \times 10^{-12}$ parts by weight. The equilibrium ratio of uranium to radium is $3{\cdot}1 \times 10^{6}$, and the proportion of uranium in the Earth's crust may therefore be estimated as 7×10^{-6}. If the whole crust had been originally composed of uranium, it would have decayed to the present proportion in seventeen half periods, or $8{\cdot}5 \times 10^{10}$ years. But this assumption would lead to the absurd conclusion that the crust must now be composed entirely of the products of radio-active disintegration of the original uranium.

The end-product of the uranium series is an isotope of lead, which appears to be stable. Now the proportion of lead in the Earth's crust is small, amounting, according to Clarke,[†] to 22×10^{-6} by weight. If this were all "uranium lead," it would result from the decay of 25×10^{-6} parts of uranium, making the initial content of the latter 32×10^{-6}, and leading to an age of 2·1 half periods, or 11×10^{9} years. Since much, if not most, of the existing lead consists of other isotopes, this estimate must be too great.

This limit may be reduced by taking account of thorium, which, according to Joly,[‡] occurs in igneous rocks in the average proportion of 2×10^{-5}. Assuming the stability of thorium lead, and taking the half period of thorium as $1{\cdot}3 \times 10^{10}$ years, it is found that the amount of lead produced by the decay of thorium during the past 8×10^{9} years, should be 10×10^{-6} of the whole mass of the crust. The decay of uranium during the same interval accounts for 12×10^{-6}, and the sum of these two reproduces Clarke's estimate of the total lead. It is of interest to note that the average atomic weight of such a mixture would be about 206·9—near that of ordinary lead, though a little low.

If uranium lead should prove to be unstable, these calculations would require revision; but the combined proportion of all the elements of atomic weight greater than 100 in the Earth's crust amounts to hardly more than one part in a thousand, according to Clarke, and almost all this is barium. On the extravagant assumption that all these elements are disintegration products, the computed age comes out $3{\cdot}5 \times 10^{10}$ years. This must be far too great. These conclusions are invalid if uranium is being formed in the Earth's crust. Building up of heavier out of lighter atoms appears to be exceedingly improbable under terrestrial conditions. The existence of a "parent" of uranium, of greater atomic weight and longer life, would meet the requirement; but such an element would have to be much more abundant than uranium, and it seems very unlikely that it would have escaped detection.

---

[*] 'Philosophical Magazine,' Series 6, vol. 24, p. 700 (1912).

[†] "Data of Geochemistry," p. 29, 'Bulletin, 695, United States Geological Survey,' Washington (1920).

[‡] *Loc. cit.*, p. 701.

There appears, therefore, to be little escape from the conclusion that the Earth's crust cannot have existed as such for so long as 3×10^{10} years, and that its age is probably much less. On the other hand, Barrell,[*] following Arthur Holmes in a discussion of the uranium-lead ratios in primary minerals, finds that minerals of Middle Pre-Cambrian Age, from localities as widely separated as Canada, Texas, Norway, and Africa, agree in indicating an age of about $1{\cdot}1 \times 10^9$ years. The age of the Earth's crust must be considerably greater.

Taking the mean of this and the upper limit found above from the ratio of uranium to lead, we obtain 4×10^9 years as a rough approximation to the age of the Earth's crust. The radio-active data alone indicate that this estimate is very unlikely to be in error in either direction by a factor as great as three. Indeed, it might be safe to say that the age of the crust is probably between two and eight thousand millions of years.

Upon the tidal (or planetesimal) theory of the origin of the Solar system (which at present appears the most probable),[†] this would be identified with the interval since the encounter which led to the ejection from the Sun of the material which formed the planets. It is then possible to postulate an origin for the terrestrial uranium, either in some heavier parent element, now confined to the denser interior portions of the Earth and Sun, or in the synthesis of heavy atoms under the extreme conditions which prevail in the Sun's interior—provided that we wish to extend our speculations so far.

It is of much interest that H. Jeffreys, from entirely different data, relating to the eccentricities of the present orbits of the planets, estimates the age of the solar system as 3×10^9 years. The agreement with the present result is far within the uncertainty of the estimates.

---

[*] 'Bulletin of the Geological Society of America,' vol. 28, p. 880 (1917).

[†] See H. Jeffreys, 'Monthly Notices of the Royal Astronomical Society,' vol. 78, pp. 424–441 (1918).

Discussion on Isotopes.

(March 3, 1921.)

Sir J. J. **Thomson**: I think I shall best open the discussion this afternoon if, instead of attempting a history of the subject of isotopes, I confine myself to a few special points. The conception of isotopes—that is, of two elements which have different atomic weights but chemical properties so closely resembling each other that they have not yet been separated by chemical methods—was, I believe, first reached by Prof. Soddy in connection with radio-active substances. In the few remarks which I have to make this afternoon, I shall not deal with these substances; I will leave them to others who are better acquainted with them than myself; I propose to confine myself to the consideration of the isotopes, or with some points connected with the isotopes, of the lighter elements.

Before doing so, I should like to say just a word about the conception of isotopes on the electron theory of an atom. On that theory the atoms of the isotopes contain an equal number of electrons. The difference in the atomic weight is supposed to be due to the simultaneous entry into the core of the atom of one or more positive charges, and an equal number of electrons, so that the electric charge on the core is not affected. In reference to this, I would like to say that it by no means follows that an electron and a positive charge will neutralise each other at the distances that occur in the atom, however close together they might be placed; and, though one would expect that the residual effect would not be large, yet I think it might easily be appreciable, and might produce some difference between the properties, chemical or physical, of the isotopes. For example, though I do not think it would affect the number of electrons in the outer layer, I think it would affect the distance of the outer layer from the core of the atom. And the statement that the chemical properties of an element depend only on the outer layer of the atom, is one with which I should agree if it is recognised that, even if the number of electrons in the outer layer is unaltered, differences in the distance of this layer from the centre may result in different chemical properties. For example, we, or at any rate some of us, believe that the valency of an element fixes the number of electrons in the outer layer. Thus, there are the same number of electrons in the outer layer of the atoms of silicon and of carbon, and yet there are well-marked differences in the chemical properties of these two substances.

With this preface, I will plunge at once into the most dramatic case of isotopes—the case of chlorine. In chlorine we have a substance with very energetic chemical properties, a substance whose atomic weight (as determined by very many observers, and with remarkably concordant results) has, I think, always been regarded as accurate to one part in a thousand, and equal to 35·5; and not only has the atomic weight of chlorine been determined by what may be called chemical methods, it has also been determined by comparing the density of the gas with the density of the standard gas, with results which were identical with those derived from chemical processes, and this fact eliminates many possible explanations which might be given for an error in the atomic weight. The position is that chemists declare the atomic weight of chlorine to be 35·5. Mr. Aston, who has measured the atomic weight of chlorine by a different method, cannot find any chlorine—or any substance—with an atomic weight of 35·5, but he does find substances with atomic weights of 35 and 37. Accepting the numbers on both sides, there does not seem to be any explanation other than that either chlorine is a mixture of these two substances of atomic weight 35 and 37, or else that, in the discharge tube which Mr. Aston employs to measure the atomic weight, some decomposition or integration—or both—of the chlorine atom has occurred.

I think the great importance of the case of chlorine is that it allows us, by an experiment of not prohibitive difficulty, to decide whether chlorine is a mixture or not, and thus settle the question of its isotopes. This is due to the fact that chlorine has such pronounced chemical properties, and that there is such a large amount, more than 30 per cent., of the isotope in the mixture. Imagine that we have a mixture of what we might call light chlorine and heavy chlorine. The molecules of the light chlorine will, on the average, have a greater velocity than those of the heavy. Thus, if we consider a reaction such as that between chlorine and arsenic, or between hydrochloric acid and an alkali, a larger proportion of the faster molecules than of the heavier ones would strike against the absorbing substances and be absorbed, so that the gas left behind would gradually increase in density. I have calculated the kind of increase that you would get from an effect of this kind. I took the case of hydrochloric acid rather than that of chlorine, because, supposing there are two chlorines, you can only get two hydrochloric acid molecules, but you could get three different molecules of chlorine; and the case is made a little simpler by dealing with hydrochloric acid rather than with chlorine. The result of my calculation is, that if you start with a given volume of HCl, and wish to make an alteration of one part in a thousand in the density, you would have to reduce the volume to 37 per cent. of its original value. If you wished to produce an alteration of ½ per cent. in the density, you would have to carry the reduction of the volume to about 1/180 of the original value, so that you would have to reduce a litre to about 5½ c.c. If you wished to produce a change of 1 per cent., you would have to reduce the volume to about 1/24,000th part of its original value, and so you would have to leave only about 0·04 c.c. of your litre. But considering the fact that the atomic weight of chlorine has been determined by many people to one part in a thousand, it would seem that it ought not to be a very difficult matter to detect a change of ½ per cent. in the atomic weight. So that it seems to me that an experiment which apparently does not present any prohibitive difficulties—though you never can be sure until you start these things—would enable one definitely to settle the question whether chlorine was a mixture or not. If I am right, there are certain processes by which one could fairly easily alter the density of chlorine. The question arises whether those chemists who have determined the atomic weight of chlorine have gone through processes anything resembling the one I have just sketched. I do not see that there is any evidence that in their determination they in any way approached to such a method of selection as that which I have suggested, so that, if the different chemists had been supplied with chlorine of the same composition, I think that they would have got by the methods they employed results for the atomic weights which were concordant to one part in a thousand. Then the question arises, Is it likely that the various samples of chlorine should be of such invariable composition? This leads us to speculations about the early history of the world, the methods in which chlorides were formed, and other subjects on which our information is not very definite; but you will notice from what I have said that though, no doubt, you might have circumstances in which chlorine of abnormal density is produced, yet whenever this abnormal chlorine is produced, a very much greater amount of normal chlorine has to be absorbed along with it. For example, chlorine which had been altered by 1 per cent. would be only 1 part in 20,000 of the chlorine in the chloride that had been produced along with it. So that the chances of getting the abnormal specimen would be somewhat small. Then again you have to remember that probably chlorides have been in solution in the sea; this would tend to mix them up; how complete the mixing would be would depend on circumstances of which we have little or no knowledge. But I do not know whether it is very profitable to speculate on this point, because I think that the whole question could be settled by a comparatively simple experiment.

I consider that it is all the more desirable that it should be settled in this way, because the information that is derived from the positive ray method gives us no direct information about the chemical properties of the substances giving the two lines corresponding to the two isotopes. Two lines may always occur together without necessarily being due to isotopes. The heavier element may be a compound of the lighter, or the lighter may be due to the dissociation of the atom of the heavier; or again it may be that the introduction of a substance like chlorine produced a kind of catalytic effect, and led to the formation of compounds which did not necessarily contain chlorine at all. The evidence as to the chemical properties which we derive from the study of the positive rays is very indirect, and it is highly desirable that some direct experiment should be made. If we take, for example, two substances found by Mr. Aston—35 and 37—we might imagine that the 37 was a compound of chlorine with two atoms of hydrogen. I know that the chemist would treat with derision a compound with that composition; so should I, if the hydride was formed under normal circumstances; but what I think is not sufficiently recognised is that the chemical properties of charged atoms—and it is charged atoms that occur in the discharge tube—are probably very different from those of uncharged ones. For example, we regard the normal chlorine atom as having seven electrons in the outer layer, and as being able, therefore, to take one hydrogen before it becomes saturated. Now if the chlorine atom becomes positively electrified, it loses an electron, and has only six in the outer layer; it is now analogous to the normal atom of oxygen, and can take up two atoms of hydrogen before being saturated, and form the compound H_2Cl. Or to take the case of oxygen, whose normal atom has six electrons in the outer layer and has room for two hydrogens; when positively electrified, it has only five electrons and has room for three hydrogens, and so can form the compound H_3O, with a molecular weight of 19. As a matter of fact, we do very often find, on the positive ray photographs, lines with an atomic weight of 19.

In passing I might mention that I think it is by no means impossible that the inert gases might, in the discharge tube, be able to form compounds. For example, neon has an outer layer of eight, but if it becomes positively electrified the eight is reduced to seven, and there is room for an atom of hydrogen, so that it is possible, though I do not say it is very probable, that neon may, in a discharge tube, form a hydride.

The behaviour of electrified atoms as distinct from normal atoms would seem to be well worthy the attention of chemists, and I think would yield interesting results. We are not helpless in the positive ray method with regard to telling whether a substance is a molecule or an atom, because there are, in general, well-marked distinctions which mark the two classes of particles. If a line is due to an atom then we generally find another line due to the atom with a double charge. It is very much rarer to find a molecule with a double charge. It is very much rarer—but that rule is not without exception—there are molecules with double charges. Another test is that when we find a positively electrified atom, in many cases we find a negatively electrified one. We do not by any means so frequently find a negatively-electrified molecule, but then again this also is a rule which is not by any means without exception. My experiments with unsaturated molecules—and I include under that term those which contain what the chemists call double bonds—show that they can, under suitable circumstances, require a double charge. Well, Mr. Aston has applied all these tests to the line corresponding to the atomic weight 37, and he finds it has a double charge; he finds also it occurs on the negative side; so it passes all tests we can apply to distinguish between a molecule and an atom, and there is thus a strong presumption that it is due to an element. But then again there is another test which I am not so sure it has passed quite so successfully. If chlorine is an invariable mixture of these two components, and if these two components are associated so closely with each other that they cannot be separated by ordinary chemical methods, then we should expect that under all circumstances the ratio of the intensities of the lines should be invariable, and should have a value which would correspond to that which would account for the abnormal atomic weight of the mixture. Some few years ago I made a large number of experiments on chlorine by the positive ray method, trying to find if the large departure of the atomic weight of chlorine from an integer was due to its being a mixture. My arrangement was by no means so accurate as Mr. Aston's, and I was not able to determine with any certainty the atomic weight, but I found the line at 37 and at first thought it was an element; the reason I gave up that assumption was that I found the proportion between the intensity of this line and that of the chlorine line was exceedingly variable. In some photographs the two were comparable in intensity; in others one was very faint compared with the other, and I could not find the 37 line at all on the negative side. So I gave up the idea that it was a separate element. I quite admit that the evidence in favour of it being an element has been raised to quite another plane by Mr. Aston's discovery of the atomic weight. That, of course, introduces—if the atomic weight is 35—a difficulty which was left entirely unexplained by my supposition about hydrides. But even yet I do not understand why I got this very large variation in the intensities, and from what I can see from Mr. Aston's paper in the 'Philosophical Magazine,' it seems to me that in the photograph there are larger variations than one would expect if the two chlorines were mixed in invariable proportion, and if their properties were practically identical. I do not mean to urge this as anything serious, but it does leave yet something to be explained, namely, how these two substances, supposed to be in invariable proportions and to have the same properties, can give rise to different ratios in the intensities of their lines.

Finally, a very interesting and important result—one which is fundamental on any theory of the ultimate constitution of matter—is that all the atomic weights which Mr. Aston has determined are integers, within the limits of his experimental error. That is a result of enormous importance. Its importance depends to a peculiar extent on the accuracy to which we can carry the determination of the atomic weight. I may perhaps illustrate my point in this way: If the atomic weight of chlorine is exactly 35, then if the atom is built up of three hydrogen atoms and eight helium ones, assuming the atomic weight of helium to be 4 and that of hydrogen 1·008, there would be a loss of mass, and therefore a liberation of energy when the chlorine atom is formed. If, on the other hand, the atomic weight of chlorine were greater than 35·024, there would be a gain of weight and an absorption of energy. Thus, an accuracy of 1 part in 2000 in the determination of the atomic weight of chlorine would be required to settle whether the formation of the chlorine atom was accompanied by an absorption or liberation of energy. Now, though I have long advocated the determination of atomic weights by positive rays, I do not think that just at present we can claim an accuracy of one part in a thousand; and, from some points of view, the focus method which Mr. Aston uses is, I think, more liable to error than the older method. One reason for this is that the quantity which Mr. Aston has to measure is the position of the edge and not the middle of a line; and one feature of the photographs of the positive rays is the way that the lines broaden when the exposure is increased. A line which looks like a spider line for a short exposure may under exactly similar circumstances become quite thick when the time of exposure is increased. Now, if you are going to measure an atomic weight to one part in a thousand, and do it step by step, starting from hydrogen, supposing the two lines corresponding to successive steps are 5 mm. apart, you will have to measure to 1/200th of a millimetre. I am very doubtful whether, when you are dealing with lines of different intensities, you can trust your measurements of the edges of these lines to anything like 1/200th part of a millimetre. In the other method, a spreading of the lines would not be so important if the thickening were symmetrical.

There is just one other point which I may illustrate by taking the case which in principle is identical with Mr. Aston's method. Suppose a horizontal beam of positive rays is deflected downwards by an electrostatic field concentrated at A, and upwards by a magnetic field concentrated at B. Then, if v is the velocity of the rays, the vertical deflection y = PN at a point, P, will be given to a first approximation by the equation

$$y = -\frac{\alpha e}{mv^2}(x+a) + \frac{\beta e}{mv}x, \tag{1}$$

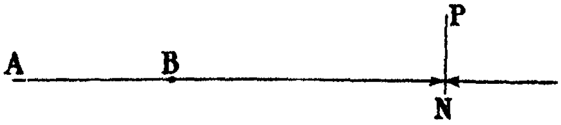


FIG. 1.

where α and β are constants depending on the electric and magnetic fields respectively; x = BN, $x+a$ = AN. The rays corresponding to different values of v form a beam, and this beam will have a bright edge, a caustic, where consecutive rays intersect. At a point on this caustic the differential of 1 with respect to v must vanish, so that

$$2\frac{\alpha e}{mv^3}(x+a) - \frac{\beta e}{mv^2}x = 0,$$

or

$$v = \frac{2\alpha}{\beta}\left(\frac{x+a}{x}\right). \tag{2}$$

Substituting this value in (1) we see that

$$y = \tfrac{1}{4}\frac{\beta^2}{\alpha}\frac{x^2}{x+a}\frac{e}{m} \tag{3}$$

is the equation to the caustic.

The caustic is thus a hyperbola passing through B; one asymptote is parallel to $y = \frac{1}{4}\frac{\beta^2 e}{\alpha m}x$, the other to $x = 0$. The particles corresponding to the different atoms and molecules in the positive rays would have different hyperbolic caustics, and if these rays fell on a photographic plate, PQ, a series of

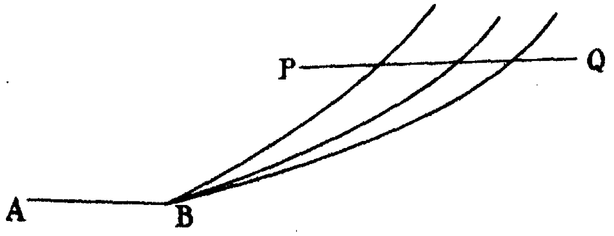


FIG. 2.

images would be formed. As, however, each point on the caustic corresponds to a special value of v, the formation of these images requires the positive rays to contain particles with the appropriate velocity. This velocity is given by equation (2), and we see that when x is large compared with a the velocity is $2\alpha/\beta$ and is the same for all particles. If the photographic plate is placed at right angles to the path of the undeflected rays, x is constant over the plate; hence by equation (2) all the rays that come to a focus on the plate must have the same velocity. If there are no particles belonging to a particular element with this velocity, the image corresponding to that element will be out of focus. We see from (3) that the deflections of the image will be inversely proportional to the atomic weight.

If the rays corresponding to each element only include a small range of velocities, then the photographic plate must be inclined so as to cut the caustics at points corresponding to those velocities. If, as in Mr. Aston's experiments, mv^2 is constant, then we have from equation (2)

$$\frac{4m\alpha^2}{\beta^2}\left(\frac{x+a}{x}\right)^2 = \mathrm{C};$$

combining this with (3) we see that

$$y = e\,\frac{\alpha}{\mathrm{C}}(x+a)$$

is the equation to the line for the photographic plate.

Thus the plate must pass through the point A. The distance, d, of the image due to a mass, m, from A, measured along the plate, is given by

$$\frac{md^2}{(d-a)^2} = \text{constant}.$$

This is far from a linear relation, so that interpolation to find the value of m for a particle giving an image between those of two particles of known masses would, without elaborate corrections, lead to results which might differ appreciably from the true values.

It would probably be only a small thing, but it would be a thing which I think would have to be taken into account in dealing with an accuracy of 1 part in 1000; but Mr. Aston has shown such great ingenuity in overcoming other difficulties that I feel no doubt that he will be successful in rendering his method free from the objection.

I have only one other word to say, and that is that this is a very striking example of the results which flow from the application of physical methods to chemistry. It is not the first case. We had the case of argon; we had the case of the structure of crystals; and I am quite sure that it will not be the last. I believe there is no branch of science where promise of great discoveries is more hopeful than those which will result by researches which involve the application of physical principles and physical measurements to chemical phenomena.

Dr. F. W. Aston: In the very short time at my command, I shall be unable to describe my apparatus, which is rather unfortunate, as several of the objections raised by the first speaker rather depend upon it. But I will deal with the points in the order in which Sir Joseph Thomson raised them. As regards the experiment suggested for the separation of hydrochloric acid, I think there is every probability of its working, if carried out under reasonable conditions. As far as I know, the mathematics of the experiment would be much the same as those in diffusions through pipeclay, the method I used in the case of the isotopes of neon, and in that case a positive separation appears to have been obtained. I obtained a difference of density —between the lightest and the heaviest sample of neon—of 0·7 per cent., which, I think, was real. As regards the addition of hydrogen to charged atoms, I presume Sir Joseph Thomson means that, if there was an atom of chlorine exactly 35·0—supposing my figures to be correct—then the other lines—the lines that one gets at 36, 37, and 38—would be due to hydrogen addition products, and that a charged chlorine might be capable of taking on one more hydrogen than an uncharged one. This is certainly a very interesting possibility, but I think it is rather discounted by the fact that if that were so, if the charged particle can take up one more hydrogen, the charged atoms of the inactive gases certainly ought to take up one hydrogen. I have seen no general indication of such hydrides, although they may be there, though too faint to be seen. But that, of course, does not get round the chemical atomic weight, because, if the fundamental chlorine is 35, it seems quite impossible to account for the chemist's atomic weight of 35·46. One must have a mixture of heavier atoms mixed with the 35 to account for that. The ratio of the intensity of the lines would be best determined by electrical methods; the photographic method is obviously a dangerous one to use for numerical values. One can certainly say that one isotope is there in greater proportion than another if its line is darker, because isotopes must have the same effect on the photographic plate; but the exact proportion is, of course, a matter that depends on the mechanism of the photography,* and therefore one cannot put too much value on the relative intensities of a plate. If one takes what seems a reasonable range of intensity, I think, with regard to the isotopes I have so far determined, their intensity corresponds very well with the actual chemical atomic weight when that is known with great certainty, as in the case of chlorine.

---

* The positive rays do not penetrate the photographic film to any appreciable extent, so that their action is not really comparable to that of light. Hence the photometric laws established for the latter will not hold in general for the purely surface effect of positive rays, though they may be approximately similar over a limited range of intensity.

[A slide was then shown of the mass spectra of chlorine, and it was indicated that a large number of lines of known mass were available with which to compare the chlorine lines. Also that the evidence was quite uncompromising both as regards the absence of a line 35·46 and the presence of a group of lines of integral mass 35, 36, 37, 38, as due to the presence of chlorine in the discharge tube.]

I have recently obtained conclusive proof from negatively charged lines, which only appear at 35 and 37, that these are the elementary lines of chlorine, 36 and 38 being hydrochloric acid.

In the matter of accuracy of measurement, I think Sir Joseph Thomson is mistaken when he says I should have to measure to 1/200th of a millimetre to get an accuracy of 1 part in 1000. The actual distance apart of two lines, say for simplicity 10 and 11, differing by 10 per cent., is about 1 cm.; hence, to compare masses to 1 part in 1000, it is only necessary to measure to one-tenth of a millimetre.

In my first experiments with neon I used the most diffracted edge of the line, as that seemed to give the most reliable results; since then the focus has been much improved, and the lines are wedge-shaped similar figures, so that one can check the results by measuring from both edges or from the middle. I admit that edge measurement should only be relied upon in the case of lines of approximately the same intensity; in that case there seems no objection to it.

[A slide was then shown of a recent analysis of rays of the element lithium, generated by means of a heated anode, which gave parabolas, clearly indicating the presence of isotopes 6 and 7.]

[Referring to a slide of the most recent mass spectra of krypton and xenon]:—Here is the xenon group, which consists of five isotopes, possibly of seven. One can see five distinctly. That is the krypton group of six lines, and here is the mercury complex, which probably consists of six lines, ending with 202 and 204. You will note that, if the hydride explanation for this group was adopted, it would be necessary for mercury to take on a very large number of hydrogen atoms, and to do so in such a way that the compounds were invariably there in the same proportion, not only in the singly charged or first order lines, but also in the second, third, fourth, and higher orders.

The lowest spectrum is that of iodine: it shows a single line at 127. Iodine is a simple element: it has no isotopes.

This is a diagram showing the relation between the isotopes of the halogen elements and those of the inactive gases. You will note that no two occupy the same number on the mass scale. If they did, they would be isobares, as described by Prof. Soddy, of the same atomic weight, but of different chemical properties. I have not found any so far, but they must exist, as otherwise there do not seem to be enough whole numbers to go round.

[Other slides of mass spectra, and a list of the isotopes so far determined, were exhibited. The principle by which rays of the desired energy were selected by a diaphragm was explained, and the close analogy with the optical spectrograph indicated. The speaker was under some misunderstanding in connection with a point raised by Sir Joseph Thomson concerning the velocity of the rays. It was a second order effect in the caustic giving the focus; it does not appear likely to affect the present measurements to 1 part in 1000.]

Prof. F. SODDY: I wish to confine my remarks mainly to the early history of this subject which had its origin, of course, in the study of radioactive change. By 1910 it had been recognised that elements identical in chemical character existed, and the question was one really of the courage of the investigator in pushing his facts to their logical conclusion. My own mind, as to the complete chemical identity of certain elements encountered in the study of radioactive change, was made up after an investigation of the chemical properties of mesothorium and radium. I submitted a mixture of these two elements to a prolonged fractional crystallisation without altering the proportions of the mixture to the smallest extent. But there were numerous previously known examples—ionium and thorium worked out by Boltwood, and radium-D, and lead—though no one probably before thought such elements were really completely identical in chemical character, but merely that they must be very much alike. An important philosophical implication followed from the view of their complete chemical identity. For if ionium and thorium are as like as, say, one piece of lead is to another piece of lead, then is one piece of lead the same as another piece of lead? What real test has the chemist of the homogeneity of the elements? None! For all the tests of the chemist are chemical and he rarely, if ever, makes practical use of such old methods as diffusion, not to speak of the new methods we have been hearing about, which would suffice to reveal differences of atomic mass in elements of identical chemical character. As far back as 1910 I summed up the situation that has been reached in my Annual Report on Radioactivity to the Chemical Society:—

"The recognition that elements of different atomic mass may possess identical chemical properties seems destined to have its most important application in the region of inactive elements, where the absence of a second radioactive nature totally unconnected with the chemical nature makes it impossible for chemical identities to be individually detected. Chemical homogeneity is no longer a guarantee that any supposed element is not a mixture of several of different atomic weights, or that any atomic weight is not merely a mean number . . . The absence of simple numerical relationships between the atomic weights becomes a matter of course rather than of surprise."

The results we have been listening to on the application of the positive rays to this question, first by Sir Joseph Thomson by the parabola method, and later, by Mr. Aston to a much greater degree of refinement, bear out the philosophical implication drawn in 1910 very thoroughly. The existence of what are now called isotopes was clearly recognised much earlier than the interpretation of them expressed by their name, which dates from 1913. The extensive researches of Dr. Fleck in Glasgow on the chemical character of all the products of radioactive change enabled the complete law of radioactive change of A. S. Russell, K. Fajans and myself to be formulated. Expulsion from the radio-element of an α-particle, carrying two positive atomic charges, causes the element to shift two places in the periodic table in the direction of diminishing mass and the expulsion of a β-ray or negative electron one place in the opposite direction. By one α- and two β-ray changes in succession the element reverts to its former place and, in this event, though the first and second products are necessarily chemically different from the parent, the third is necessarily identical with the parent. Moreover, the means of detection in this case gives the means of separation. For, if the sequence of one α- and two β-changes is represented $A \rightarrow B \rightarrow C \rightarrow D$, where where A and D are necessarily chemically identical and different from B and C, B and C may be separated chemically from A and D, and in course of time will generate D free from A, so that D can by a second separation be obtained alone. This double separation, a suitable interval elapsing between the two separations, is a common method of separating isotopes genetically connected. For example, radiothorium free from its isotope thorium is so obtained commercially from thorium minerals by separating the mesothorium and leaving it to grow fresh radiothorium after separation. There is no other way of separating it.

Writing a book in 1913 on the chemistry of the radio-elements I got tired of writing "elements chemically identical and non-separable by chemical methods" and coined the name *isotope*, signifying the same place, to express their nature in the light of the new interpretation of the periodic law. It was clear that the "places" of the latter expressed unit difference of charge in the atomic constitution, a conception which the researches of Moseley, on the wave-lengths of characteristic X-rays, later extended throughout the whole table. The nuclear theory of atomic structure of Sir Ernest Rutherford dovetailed in nicely. The α- and β-rays being expelled from the central nucleus of the atom, the expulsion of one α- and two β-rays leaves the *nett* nuclear charge and the whole external electronic system unchanged. Isotopes are, thus, elements with atoms made up of identical external systems and nuclei with the same *nett* charge. The constitution of the nucleus and the absolute number of contained + and − charges are different, but these differences are without effect on the chemical character.

The correctness of these views received very powerful support from the study of the atomic weight of lead derived from radioactive minerals. Both the uranium and thorium disintegrations end with the place in the periodic table occupied by lead, but whereas common lead has an atomic weight 207·2, that from uranium (atomic weight 238), formed by the successive expulsion of eight α-particles or helium atoms, should be 206 and that for thorium (atomic weight 232) by five α-particles, should be 208. The atomic weight of lead from the purest uranium minerals has been found to be 206·05 and from the purest thorium minerals 207·9. Moreover, thorium separated from uranium minerals contains the isotope ionium, 2 units less in atomic weight. Hönigschmid found, for a thorium preparation weighing a few grammes separated from 30 tons of Joachimsthal pitchblende by von Welsbach, an atomic weight 231·51 as compared with 232·12 for a pure preparation of ordinary thorium. The spectra of the two were identical and showed a complete absence of all impurities. This is one of the finest examples of the extent to which the complete chemical identity of isotopes has been experimentally established.

The results we have been listening to have shown what a powerful new weapon has been developed for the detection of isotopes among common elements not in actual course of change. Sir Joseph Thomson alluded to chlorine as the most dramatic example of isotopes, but, personally, I consider the discovery of metaneon, an isotope of neon, by himself and Mr. Aston, deserves that title. Coming in 1913, just as the theory of the radioactive isotopes had been eludicated, it constituted a most striking and suggestive confirmation, in a region of the periodic table remote from that in which the conception originated.

With regard to the interesting point raised by Sir Joseph Thomson, that the diameter of the outer atomic shell, though not the number of the electrons contained therein, may be affected by the gross charge of the nucleus, the nett charge being the same, so far as we can tell it is not. There is no trace of such an effect as you have—as in the example to which he referred, carbon and silicon—between two members of the same family. The atomic volumes of the two kinds of lead referred to are the same. I actually determined the atomic weight of my thorium lead from the density, on the assumption that the atomic volume was the same as that of ordinary lead, and got the value 207·74. A single chemical determination gave the value 207·69, and, in a more complete investigation, Hönigschmid, in Vienna, with another fraction of the same lead, found 207·77 ($\pm 0·014$). T. W. Richards and Wadsworth have since found that the atomic volume of uranium lead of atomic weight 206·085 is the same as that of ordinary lead.

As regards the method suggested by Sir Joseph Thomson for the separation of the isotopes of chlorine by allowing hydrochloric acid to be absorbed by alkali, as a chemist I would differ from Mr. Aston's opinion that it would be analogous to the diffusion method which he has used in an attempt to separate the isotopes of neon, because, as a matter of fact, we have no conception as to why one element combines with another. It certainly has little to do with molecular speed, as the temperature coefficient of chemical change shows.

Sir JOSEPH THOMSON: I think you misunderstood me. I did not say there was any peculiarity in the affinity, so to speak, between the surface and the chemical. But the mere fact that some molecules were moving faster would cause them to strike oftener.

Prof. SODDY: Then this brings it down to the ordinary diffusion effect, which has been tried for neon and metaneon with somewhat divergent results.

I think Mr. Aston is to be very much congratulated upon the outcome of what is probably one of the most brilliant combinations of mathematical analysis and experimental skill this century has produced. I do not think chemists have any reason to doubt the accuracy, to one part in a thousand, which he claims for his determinations. He has on his plates a great range of thoroughly standard lines on which to test the accuracy, and I am prepared to accept his results as of the same order of approximation as is attained in the best chemical atomic weight methods.

Isotopes have been of the utmost use to chemists in the laboratory in various ways. Their use as indicators, developed by Paneth and von Hevesey, has enabled, for example, the solubility of very sparingly soluble salts of lead to be determined with ease. By their aid it has been possible to test the Nernst theory of E.M.F. at concentrations of the order of 10^{-9}, completely beyond the range of ordinary methods. Recently they have given practically a direct demonstration of the separated existence of the ions in an electrolyte. Thus, if ordinary lead chloride and a lead nitrate that has been "activated" by addition of thorium-B, an isotope of lead, are mixed in pyridine solution, the lead chloride crystallising is half as active as the original lead nitrate. But if, instead of being mixed with a salt that is ionised in solution, an active lead salt is mixed with inactive lead tetraphenyl, in a suitable solvent, the latter is found to be inactive when again separated. By the use of the same isotope the rate of self-diffusion of lead molecules through lead molecules has recently been experimentally measured.

Lastly, the existence of isotopes seems to offer in some cases a searching test of chemical logic,—but I have already taken more than I should of the time of the meeting.

Prof. T. R. MERTON: The most recent results of Dr. Aston and Mr. Thomson relating to the two isotopes of lithium provide a case in which a spectroscopic test can be applied more easily than in the case of any other isotopes. It is one of the greatest successes of Bohr's theory that the difference between the wave-lengths of the hydrogen lines of the Balmer series and alternate members of the ζ Puppis series of helium can be accounted for by the difference in the masses of the nuclei, and this can be applied in the same way to the two isotopes of lithium, of atomic weight 6 and 7. The structure of the red lithium line of wave-length 6708 A. has been investigated by several observers, and under normal circumstances consists of a doublet with a difference in wave-length of 0·151 A., the violet component being the stronger member. Now the doublet due to one isotope should be separated from the doublet due to the other by a difference in wave-length which can be calculated on exactly the same basis as in the case of the Balmer hydrogen lines and the lines of the ζ Puppis series; and if that is done, we find that the one doublet is displaced from the other by 0·087 A., and from the accepted atomic weight of lithium, 6·94, one doublet would presumably be about one-sixteenth of the intensity of the other. This is, from a photometric point of view, not a very great difference in intensity, three magnitudes only, and the displaced component should easily be visible. I am re-examining the lithium spectrum under a very high dispersion, and though the results are not yet complete, I have so far not been able to find any evidence of these faint components, and I think that if they are present at all, they must be very much less than one-sixteenth of the intensity of the main doublet. Some time ago, Prof. Lindemann and I tried experiments of the same kind with neon; in this case the difference between the two components should be much smaller, and on account of the physical widths of the lines, it was impossible to reach any conclusion, but the experiment will be repeated at the temperature of liquid hydrogen, and this should yield definite results.

A short time ago, General Hartley and I suggested a photochemical method of separating the isotopes of chlorine, but this method has not so far been successful. I have made some experiments on a method of separating possible isotopes of zinc. If zinc consists of two isotopes, as presumably it does, from the fractional value of its atomic weight, and if there is even a very small difference in the electrical properties of these two isotopes, it seemed possible that if a small quantity of zinc were kept fused for a long time under a large quantity of fused zinc chloride, there might be a settling-out of one of the isotopes into the small quantity of zinc; and although this investigation is not really finished, neither the densities nor the atomic weights having been determined, the spectra of the zinc treated in this way and of ordinary zinc appear to be identical to a very high degree of accuracy. The results for lead and thallium have already been published, and they show a perfectly definite difference in wave-length, which is in very close agreement in the case of uranium lead with the results obtained by Harkins and Aronberg in America. These differences are greater than would be expected from theory, but the lead atom is presumably very complex, and no definite conclusion can at present be drawn from this result. The case of lithium appears to be more crucial.

Prof. F. A. LINDEMANN: The only point which has been raised so far to which I cannot agree is the statement that isotopes have identical chemical properties. That they are very similar is indubitable, but I think it may be shown generally that two substances of different atomic weight must be separable chemically in principle.

If K_p is the equilibrium constant in the law of mass action it may be shown that $\frac{\log K_{p_1}}{\log K_{p_2}} = RT \log \frac{p_2}{p_1}$. If the force acting upon an atom in a solid is $\phi(r)$, r being the distance of the neighbouring atoms,

$$\log p = + \int_r^{\infty} \frac{\phi(r)\,dr}{K_1 T} + \int^T \frac{dT}{K_2 T^2} \int^T \left[K_3 - f\left(\frac{\phi'(r)}{M}\right)\right] dT + i,$$

and since K_1, K_2, and K_3 are constants, K_{p_1} can only be equal to K_{p_2} at all temperatures, *i.e.*, the substance can only be inseparable, if

$$f\left(\frac{\phi'(r_1)}{M_1}\right) \equiv f\left(\frac{\phi'(r_2)}{M_2}\right) \text{ and } \int_{r_1}^{\infty} \phi(r)\,dr \equiv \int_{r_2}^{\infty} \phi(r)\,dr.$$

Since $r_1 = r_2$, these equations cannot be rigorously fulfilled unless $M_1 = M_2$. The difference will be very small if the chemical constants are given by $\frac{3}{2} \log M -$constant, and if there is a Nullpunktsenergie. The failure so far to separate isotopes seems to me to be an argument in favour of these two assumptions.

It seems highly probable that isotopes of diatomic molecules such as Cl_2 should have different spectra. According to the usual view the frequency in band spectra is given by $\nu = \nu_0 + \frac{h}{8\pi^2 I}(n_1^2 - n_2^2) = \nu_0 + \frac{h}{8\pi^2 I}(\pm 2n+1)$ since $n_1 - n_2 = \pm 1$. Here the moment of inertia I is that of the molecule, *i.e.*, $I = \frac{M_1 M_2}{M_1 + M_2} r^2$, since the distance apart r is constant. In Cl_2 if M_1 and M_2 are thirty-five and thirty-seven times M_0, therefore the moment of inertia can have any of the three values

$$r^2 M_0 \frac{35}{2}, \qquad r^2 M_0 \frac{1295}{72} \quad \text{or} \quad r^2 M_0 \frac{37}{2}.$$

Hence any line should be accompanied by two others, the three frequencies being given by

$$\nu_1 = \nu_0 + (\pm 2n+1) \frac{h}{8\pi^2 r^2 M_0} \frac{2}{35},$$

$$\nu_2 = \nu_0 + (\pm 2n+1) \frac{h}{8\pi^2 r^2 M_0} \frac{72}{1295},$$

and

$$\nu_3 = \nu_0 + (\pm 2n+1) \frac{h}{8\pi^2 r^2 M_0} \frac{2}{37}.$$

Since $M_0 = 1{\cdot}65 \,.\, 10^{-24}$ and $r_0 < 3{\cdot}10^{-8}$, $\Delta\nu > (2n \pm 1) \,.\, 8{\cdot}6 \,.\, 10^7$, which for 5000 Å. amounts to $\Delta\lambda > 7{\cdot}10^{-4}(2n \pm 1)$ Å. In other words, the chlorine spectrum should consist of numerous triplets of constant intensity ratio.

It might be possible to separate *isotopes electrolytically if*, as seems likely, they have different migration velocities in solution. If an ion of mass M retains the *fraction* $\frac{M}{M+m}$ *of its momentum on collision with* a solvent molecule of mass m, the velocity component due to the field is given by

$$\left(V_1 + \frac{Ft}{M}\right) \frac{M}{M+m} = V_1 \qquad \text{or} \qquad V_1 = \frac{Ft}{m},$$

V_1 being the remnant velocity at the beginning of a free path, F the electric force, and t the time between collisions. The mean velocity of migration is therefore

$$V = V_1 + \frac{Ft}{2M} = Ft\,(1/m + 1/2M).$$

If the free path l is the same for different isotopes

$$t = \frac{l}{V_0} = \frac{l}{\sqrt{(3kT)}} \sqrt{(M)},$$

V_0 being the equipartition velocity of thermal agitation.

Hence the migration velocity

$$v = \frac{l}{2m\sqrt{(3kT)}} \cdot \frac{2M+m}{\sqrt{(M)}} \quad \text{and} \quad 1/v \, \frac{dv}{dm} = \frac{1}{2M} \cdot \frac{2M-m}{2M+m}.$$

For lithium, where M_1 and M_2 are 6 and 7 if $m = 18$, there should be a difference of 1·24 per cent. in the migration velocity. For chlorine with $M_1 = 35$ and $M_2 = 37$, it should be nearly 1·7 per cent. If a solution were electrolysed in some medium like gelatine, the head of the column should contain one isotope only, and it is hoped shortly to carry out experiments to test this. In the above calculation, of course, the effect of hydration and association have not been taken into account.

One other point to which it may be worth while to refer is the very small energy difference between atoms if their atomic weights approach whole numbers closely. Quite apart from positive rays, chemical evidence shows that the difference between the mass of one carbon atom and three helium atoms is so small that the energy difference cannot exceed $4{\cdot}5 \,.\, 10^{-7}$ ergs per atom, *i.e.*, about 1/30 of the energy of the average radioactive change. If this is true, it somewhat curtails the store of intra-atomic energy, which it is hoped will one day become available.

A Generalisation of Weyl's Theory of the Electromagnetic and Gravitational Fields.

By A. S. Eddington, F.R.S.

(Received February 19, 1921.)

1. In the non-Euclidean geometry of Riemann, the metric is defined by certain quantities, $g_{\mu\nu}$, which are identified by Einstein with the potentials of the gravitational field. H. Weyl has shown that, on removing a rather artificial restriction in Riemann's geometry, the expression for the metric includes also terms which are identified with the four potentials of the electromagnetic field. I believe that Weyl's geometry, far-reaching though it is, yet suffers from an unnecessary and harmful restriction; and it is the object of this paper to develop a still more general theory.

In passing beyond Euclidean geometry, gravitation makes its appearance; in passing beyond Riemannian geometry, electromagnetic force appears; what remains to be gained by further generalisation? Clearly, the non-Maxwellian binding-forces which hold together an electron. But the problem of the electron must be difficult, and I cannot say whether the present generalisation succeeds in providing the materials for its solution. The present paper does not seek these unknown laws, but aims at consolidating the known laws. I hope to show that, in freeing Weyl's geometry from its limitation, the whole scheme becomes simplified, and new light is thrown on the origin of the fundamental laws of physics.

The essential point of Weyl's theory† is that comparisons of length at different times and places may yield discordant results according to the route of comparison. Accordingly, a particular standard of length should only be used at the time and place where it is; and it is necessary to set up a separate unit of length (interval-length) at every point of space and time. Such a system of unit-standards is called a gauge-system, and Weyl's theory proceeds on the view that the gauge-system is arbitrary, just as the coordinate system is. Nevertheless, he recognises that a certain natural gauge-system plays a very important part in physics. We *do* compare lengths on the sun and the earth: and, although our method of comparison may be ambiguous for the thirtieth significant figure, we have no hesitation in deciding whether a given length on the sun should be called an inch or a mile. This natural gauge-system seems at first like an excrescence on the purely geometrical theory, diverting it from its natural development, but necessary in order to bring it into touch with physics. The theory here developed seems to throw new light on the meaning of this natural gauge-system.

A deductive study of world-geometry and an inductive study of observational science attack the problem of Nature from opposite ends. The most elementary concepts of experimental knowledge may not appear until a late stage in the deductive treatment. Our work thus falls into two stages: first, the development of a pure geometry of a very general kind; second, a physical theory based on the identification of the geometrical functions with quantities obtained by experimental measurement. The introduction of the natural gauge-system marks the transition from pure geometry to physics.

Geometrical Theory.

2. Some of the tensors used in the analysis possess the property of being unchanged by any alteration of gauge-system, in addition to their ordinary properties with regard to transformation of co-ordinates. I distinguish these by the prefix "in-" (in-tensor, in-vector, in-invariant). In symbolic notation, a star (*) indicates quantities generalised from the corresponding quantities

---

† 'Berlin. Sitzungsberichte,' May 30, 1918. Further developments are scattered throughout his book 'Raum, Zeit, Materie.' See also his article, 'Nature,' February 17, 1921. I need scarcely say that I am under great obligations to these sources.

in Riemannian geometry so as to maintain this "in-" property, wholly or partly.

A displacement, dx_μ, is a simple example of an in-vector. Its components are differences of coordinates (pure numbers), and have no reference to any gauge. Since dx_μ is contravariant, we often write it with the notation $(dx)^\mu$, which reminds us of its character.

Let A^μ be an in-vector representing an infinitesimal displacement at a point P. Can we find at an infinitely near point, P′, an exactly equivalent displacement? Clearly, the equivalent displacement will not in general be represented by A^μ, because, for example, in a *uniform* field of force the polar components vary from point to point. There will be a spurious change of A^μ, attributable to the curvilinearity of the coordinates, even when the in-vector is displaced "without absolute change." The most general possible linear expression for this spurious change, when the in-vector is moved through a displacement, dx_ν, is

$$-\Gamma_{\nu\alpha}^{\mu} A^\alpha (dx)^\nu, \tag{2·1}$$

where $\Gamma_{\nu\alpha}^{\mu}$, which is not assumed to be a tensor, represents arbitrary coefficients, and the expression is summed according to the usual convention. Only linear terms need be taken, because A^α and $(dx)^\nu$ are both infinitesimals.

Following Levi-Civita, displacement "without absolute change" is called *parallel displacement.* For parallel displacement, the change dA^μ must be equal to (2·1); and accordingly

$$\frac{\partial A^\mu}{\partial x_\nu} + \Gamma_{\nu\alpha}^{\mu} A^\alpha = 0, \tag{2·2}$$

is the condition for parallel displacement.

The definiteness of the notion of parallel displacement depends on the recognition of some kind of equivalence between displacements at infinitely near points—not between the analytical expressions, but between the *relations* which they express. We assume that there is such an equivalence; and, in justification, we point out that it seems to be the minimum assumption possible. For if there were no comparability of relations, even the most closely adjacent, the continuum would be divested of even the rudiments of structure, and nothing in nature could resemble anything else.

If in (2·1) we take A^α to be a displacement $PP_1 = (\delta x)^\alpha$, which on parallel displacement to P′ becomes $P'P_1'$, the difference of co-ordinates of P_1' and P will be

$$(dx)^\mu + (\delta x)^\mu - \Gamma_{\nu\alpha}^{\mu} (\delta x)^\alpha (dx)^\nu. \tag{2·3}$$

Interchanging the two displacements, *i.e.*, displacing PP′ along PP_1, we shall not arrive at the same point P_1', unless

$$\Gamma_{\nu\alpha}^{\mu} = \Gamma_{\alpha\nu}^{\mu}. \tag{2·4}$$

This is a necessary condition for what is called *affine* geometry (fully treated in Weyl's 'Raum, Zeit, Materie'). It appears to express the condition that the world is "flat in its smallest parts" or that it possesses a definite tangent. I shall adopt the condition (2·4) in the present investigation, ruling out an "infinitely crinkled" world.†

3. Let an in-vector A^μ be carried by parallel displacement round a small complete circuit. The difference of the final and initial values is

$$
\begin{aligned}
[A^\mu]^{\text{final}}_{\text{initial}} &= \int \frac{\partial A^\mu}{\partial x_\nu}\, dx_\nu \\
&= -\int \Gamma_{\nu\alpha}{}^{\mu} A^\alpha (dx)^\nu \text{ by } (2{\cdot}2) \\
&= -\tfrac{1}{2} \iint \left\{ \frac{\partial}{\partial x_\sigma} (\Gamma_{\nu\alpha}{}^{\mu} A^\alpha) - \frac{\partial}{\partial x_\nu} (\Gamma_{\sigma\alpha}{}^{\mu} A^\alpha) \right\} dS^{\nu\sigma}, \qquad (3{\cdot}1)
\end{aligned}
$$

by Stokes's theorem. The double integral is taken over any surface bounded by the circuit, and $dS^{\nu\sigma}$ is the antisymmetrical tensor representing the element of surface. The factor $\frac{1}{2}$ is needed because each component of surface appears twice in the summation, *e.g.*, as dS^{12} and $-dS^{21}$.

The integrand is equal to

$$
\begin{aligned}
&A^\alpha \left(\frac{\partial}{\partial x_\sigma} \Gamma_{\nu\alpha}{}^{\mu} - \frac{\partial}{\partial x_\nu} \Gamma_{\sigma\alpha}{}^{\mu} \right) + \Gamma_{\nu\alpha}{}^{\mu} \frac{\partial A^\alpha}{\partial x_\sigma} - \Gamma_{\sigma\alpha}{}^{\mu} \frac{\partial A^\alpha}{\partial x_\nu} \\
&\qquad = A^\alpha \left(\frac{\partial}{\partial x_\sigma} \Gamma_{\nu\alpha}{}^{\mu} - \frac{\partial}{\partial x_\nu} \Gamma_{\sigma\alpha}{}^{\mu} \right) - \Gamma_{\nu\alpha}{}^{\mu} \Gamma_{\sigma\rho}{}^{\alpha} A^\rho + \Gamma_{\sigma\alpha}{}^{\mu} \Gamma_{\nu\rho}{}^{\alpha} A^\rho \text{ by } (2{\cdot}2).
\end{aligned}
$$

Changing the dummy suffix from α to ρ in the first term, this becomes

$$-{}^*B_{\rho\nu\sigma}{}^{\mu} A^\rho,$$

where
$$ {}^*B_{\rho\nu\sigma}{}^{\mu} = -\frac{\partial}{\partial x_\sigma} \Gamma_{\nu\rho}{}^{\mu} + \frac{\partial}{\partial x_\nu} \Gamma_{\sigma\rho}{}^{\mu} + \Gamma_{\nu\alpha}{}^{\mu} \Gamma_{\sigma\rho}{}^{\alpha} - \Gamma_{\sigma\alpha}{}^{\mu} \Gamma_{\nu\rho}{}^{\alpha}. \qquad (3{\cdot}2)$$

Hence
$$[A^\mu] = \tfrac{1}{2} \iint {}^*B_{\rho\nu\sigma}{}^{\mu} A^\rho dS^{\nu\sigma}. \qquad (3{\cdot}3)$$

In Einstein's theory, even without Weyl's generalisation, the integrand does not usually vanish, and the initial and final values of A^μ are unequal. Of course, the same discrepancy appears if A^μ is displaced by different routes to another point, and the difference is shown by (3·3) to be proportional to the area of the circuit, *ceteris paribus*. The formula (3·3) applies only to small circuits for which the square of the area can be neglected, because A^ρ is ambiguous except on the boundary of the circuit where the route of transfer has been defined; but the ambiguity of A^ρ is of the same order of magnitude as $[A^\mu]$, and its surface integral becomes negligible compared with $[A^\mu]$ when the circuit is sufficiently small. Since then the circuit is infinitesimal, we can omit the integral sign, obtaining

---

† A coordinate-system enumerates the points of a continuum in particular *order* and this order is unchanged by any transformation of coordinates. If the world referred to an arbitrary coordinate-system turns out to be infinitely crinkled, I think this must indicate a disharmony between the coordinate-order and the structural order of the points. The remedy is to change the coordinate enumeration, not by continuous transformation, but by re-sorting the points. Evidently our displacement dx_μ will not be the measure of any physical (structural) relation unless some preliminary agreement of order is postulated. The problem of determining the conditions of this preliminary agreement is essentially that solved by A. A. Robb in his 'Theory of Time and Space.' If ever the theory comes to deal with essentially discontinuous phenomena (*e.g.*, quanta) the condition (2·4) may have to be reconsidered.

$$[A^\mu] = \tfrac{1}{2}\,{}^*B_{\rho\nu\sigma}{}^{\mu} A^\rho dS^{\nu\sigma}. \tag{3·4}$$

Now $[A^\mu]$ is the difference of two vectors at the same point, and is therefore a vector; $dS^{\nu\sigma}$ is a tensor; hence ${}^*B_{\rho\nu\sigma}{}^{\mu}$ is a tensor. Moreover, it is an in-tensor, since there has been as yet no reference to any gauge.

We obtain another in-tensor by setting $\mu = \sigma$, *i.e.*, contracting,

$${}^*G_{\rho\nu} = {}^*B_{\rho\nu\sigma}{}^{\sigma} = -\frac{\partial}{\partial x_\sigma}\Gamma_{\nu\rho}{}^{\sigma} + \frac{\partial}{\partial x_\nu}\Gamma_{\sigma\rho}{}^{\sigma} + \Gamma_{\nu\alpha}{}^{\sigma}\Gamma_{\sigma\rho}{}^{\alpha} - \Gamma_{\sigma\alpha}{}^{\sigma}\Gamma_{\nu\rho}{}^{\alpha}. \tag{3·5}$$

These two in-tensors express intrinsic properties of the continuum. So far as I can see no others (except products of these) exist. The first is a description as complete as possible of the structure of the continuum (the inter-connection of relations) so far as it can be studied by these methods; the second is an abbreviated summary of the information contained in the first.

We shall find later that physics has in the main contented itself with studying the abridged edition of the book of nature; only in the problems of electron structure does it seem necessary to search out the more extensive information contained in the original ${}^*B_{\rho\nu\sigma}{}^{\mu}$.

4. In order to introduce the $g_{\mu\nu}$ we must adopt a definite, but at present arbitrary, system of gauges. The length (interval length) l of a displacement A^μ is defined by

$$l^2 = g_{\mu\nu}A^\mu A^\nu, \tag{4·1}$$

in agreement with Einstein's fundamental equation

$$ds^2 = g_{\mu\nu}dx_\mu dx_\nu.$$

We take $g_{\mu\nu}$ to be a tensor in order to make l^2 an invariant. The tensor $g_{\mu\nu}$ may be selected quite arbitrarily,† and then equation (4·1) shows how a length, l, is to be assigned to any displacement, dx_μ, given in coordinate-differences. Different choices of $g_{\mu\nu}$ will give different plans of assigning length—in short, different gauge-systems. Length here is still a geometrical convention rather than a physical conception. So long as we have no independent means of comparing lengths, no inconsistency can possibly be detected in our scheme. When we come to identify these geometrical lengths with physical lengths, which can be compared by experimental methods, conditions must be introduced which will limit our choice of $g_{\mu\nu}$. But that will come later; the present analysis applies to any gauge-system whatever.

---

† For convenience, however, we limit the choice to tensors symmetrical in μ and ν.

Move the displacement A^μ by parallel displacement through dx_σ. Differentiating (4·1) we have

$$\begin{aligned} d(l^2) &= \left(\frac{\partial g_{\mu\nu}}{\partial x_\sigma} A^\mu A^\nu + g_{\mu\nu} A^\nu \frac{\partial A^\mu}{\partial x_\sigma} + g_{\mu\nu} A^\mu \frac{\partial A^\nu}{\partial x_\sigma}\right) dx_\sigma \\ &= \left(\frac{\partial g_{\mu\nu}}{\partial x_\sigma} A^\mu A^\nu - g_{\mu\nu} \Gamma_{\sigma\alpha}{}^\mu A^\nu A^\alpha - g_{\mu\nu} \Gamma_{\sigma\alpha}{}^\nu A^\mu A^\alpha\right) dx_\sigma \text{ by } (2{\cdot}2) \\ &= \left(\frac{\partial g_{\mu\nu}}{\partial x_\sigma} - g_{\alpha\nu} \Gamma_{\sigma\mu}{}^\alpha - g_{\mu\alpha} \Gamma_{\sigma\nu}{}^\alpha\right) A^\mu A^\nu (dx)^\sigma, \end{aligned}$$

by interchanging dummy suffixes.

In conformity with the usual rule for lowering suffixes, we write

$$\Gamma_{\sigma\mu,\nu} = g_{\alpha\nu} \Gamma_{\sigma\mu}{}^\alpha,$$

so that

$$d(l^2) = \left(\frac{\partial g_{\mu\nu}}{\partial x_\sigma} - \Gamma_{\sigma\mu,\nu} - \Gamma_{\sigma\nu,\mu}\right) A^\mu A^\nu (dx)^\sigma. \tag{4·2}$$

But $d(l^2)$, the difference of two invariants, is an invariant. Hence the quantity in the bracket is a tensor. We denote it by $2K_{\mu\nu,\sigma}$. Thus

$$2K_{\mu\nu,\sigma} = \frac{\partial g_{\mu\nu}}{\partial x_\sigma} - \Gamma_{\sigma\mu,\nu} - \Gamma_{\sigma\nu,\mu}. \tag{4·3}$$

Similarly

$$2K_{\mu\sigma,\nu} = \frac{\partial g_{\mu\sigma}}{\partial x_\nu} - \Gamma_{\nu\mu,\sigma} - \Gamma_{\nu\sigma,\mu},$$

$$2K_{\nu\sigma,\mu} = \frac{\partial g_{\nu\sigma}}{\partial x_\mu} - \Gamma_{\mu\nu,\sigma} - \Gamma_{\mu\sigma,\nu}.$$

Adding these and subtracting (4·3), we have

$$\Gamma_{\mu\nu,\sigma} = \tfrac{1}{2}\left(\frac{\partial g_{\mu\sigma}}{\partial x_\nu} + \frac{\partial g_{\nu\sigma}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\sigma}\right) - K_{\mu\sigma,\nu} - K_{\nu\sigma,\mu} + K_{\mu\nu,\sigma}. \tag{4·4}$$

We write

$$S_{\mu\nu,\sigma} = K_{\mu\nu,\sigma} - K_{\mu\sigma,\nu} - K_{\nu\sigma,\mu}, \tag{4·5}$$

where $S_{\mu\nu,\sigma}$ is evidently a tensor symmetrical in μ and ν. Then, raising the suffixes by multiplying through by $g^{\sigma\tau}$, (4·4) becomes

$$\Gamma_{\mu\nu}{}^\tau = \{\mu\nu,\tau\} + S_{\mu\nu}{}^\tau. \tag{4·6}$$

We see that $\Gamma_{\mu\nu}{}^\tau$ is a generalisation of the Christoffel symbol, and might appropriately be denoted by $^*\{\mu\nu,\tau\}$. Though not a tensor, it possesses the "in-" property, since it was introduced prior to any gauge system. We note that $K_{\mu\nu,\sigma}$ is not an in-tensor, since $d(l^2)$ depends on the gauge.

It is here that we make an important divergence from Weyl's geometry; his theory is obtained by giving to $K_{\mu\nu,\sigma}$ the special form $g_{\mu\nu}\phi_\sigma$.

5. From (3·2) and (4·6) we can evaluate $^*B_{\mu\nu\sigma}{}^\rho$. If the terms in $S_{\mu\nu}{}^\sigma$ are dropped, we obtain immediately the usual expression for the Riemann-Christoffel tensor $B_{\mu\nu\sigma}{}^\rho$ [Report,† equation (23)]. The additional terms are

$$\frac{\partial}{\partial x_\nu} S_{\mu\sigma}{}^\rho - \frac{\partial}{\partial x_\sigma} S_{\mu\nu}{}^\rho + \{\mu\sigma, \alpha\} S_{\nu\alpha}{}^\rho + \{\nu\alpha, \rho\} S_{\mu\sigma}{}^\alpha - \{\mu\nu, \alpha\} S_{\sigma\alpha}{}^\rho$$
$$- \{\sigma\alpha, \rho\} S_{\mu\nu}{}^\alpha + S_{\nu\alpha}{}^\rho S_{\mu\sigma}{}^\alpha - S_{\sigma\alpha}{}^\rho S_{\mu\nu}{}^\alpha.$$

The first six terms easily reduce to

$$(S_{\mu\sigma}{}^\rho)_\nu - (S_{\mu\nu}{}^\rho)_\sigma,$$

where the final suffix denotes covariant differentiation according to the ordinary rule, viz.,

$$(S_{\mu\sigma}{}^\rho)_\nu = \frac{\partial}{\partial x_\nu} S_{\mu\sigma}{}^\rho - \{\mu\nu, \alpha\} S_{\alpha\sigma}{}^\rho - \{\sigma\nu, \alpha\} S_{\mu\alpha}{}^\rho + \{\alpha\nu, \rho\} S_{\mu\sigma}{}^\alpha.$$

Hence

$$^*B_{\mu\nu\sigma}{}^\rho = B_{\mu\nu\sigma}{}^\rho + (S_{\mu\sigma}{}^\rho)_\nu - (S_{\mu\nu}{}^\rho)_\sigma + S_{\nu\alpha}{}^\rho S_{\mu\sigma}{}^\alpha - S_{\sigma\alpha}{}^\rho S_{\mu\nu}{}^\alpha. \qquad (5·1)$$

Setting $\rho = \sigma$, we have

$$^*G_{\mu\nu} = G_{\mu\nu} + 2\kappa_{\mu\nu} - (S_{\mu\nu}{}^\sigma)_\sigma + S_{\alpha\nu}{}^\beta S_{\beta\mu}{}^\alpha - 2\kappa_\alpha S_{\mu\nu}{}^\alpha, \qquad (5·2)$$

where we have written

$$2\kappa_\mu = S_{\sigma\mu}{}^\sigma, \qquad (5·3)$$

and $\kappa_{\mu\nu}$ is the covariant derivative of κ_μ. Again, multiplying by $g^{\mu\nu}$

$$^*G = G + 2\kappa^\nu{}_\nu + \lambda^\nu{}_\nu + S_{\nu\alpha,\beta} S^{\nu\beta,\alpha} + 2\kappa_\alpha \lambda^\alpha \qquad (5·4)$$

where

$$\lambda_\mu = -S^\sigma{}_{\sigma,\mu}. \qquad (5·5)$$

(The distinction between (5·3) and (5·5) is that in $S_{\sigma\mu}{}^\sigma$ μ occupies one of the two symmetrical places, and in $S^\sigma{}_{\sigma,\mu}$ it occupies the third unsymmetrical place. κ_μ and λ_μ are quite different vectors.)

The only term on the right of (5·2) which is not symmetrical in μ and ν is $2\kappa_{\mu\nu}$, which may be written $(\kappa_{\mu\nu} + \kappa_{\nu\mu}) + (\kappa_{\mu\nu} - \kappa_{\nu\mu})$. Thus, if we write

$$^*G_{\mu\nu} = R_{\mu\nu} + F_{\mu\nu}, \qquad (5·6)$$

where $R_{\mu\nu}$ is symmetrical and $F_{\mu\nu}$ antisymmetrical, we have

$$F_{\mu\nu} = \kappa_{\mu\nu} - \kappa_{\nu\mu}$$
$$= \frac{\partial \kappa_\mu}{\partial x_\nu} - \frac{\partial \kappa_\nu}{\partial x_\mu}, \qquad (5·7)$$

the covariant curl being the same as the ordinary curl for a *covariant* vector.

---

† 'Report on the Relativity Theory of Gravitation,' Physical Society (Fleetway Press).

Since $R_{\mu\nu}$ is half the sum of $^*G_{\mu\nu}$ and $^*G_{\nu\mu}$, and $F_{\mu\nu}$ is half the difference, both $R_{\mu\nu}$ and $F_{\mu\nu}$ are in-tensors.[†]

Similarly we can divide $^*B_{\mu\nu\sigma\rho}$ into two parts

$$^*B_{\mu\nu\sigma\rho} = R_{\mu\nu\sigma\rho} + F_{\mu\nu\sigma\rho}, \tag{5·8}$$

where R is antisymmetrical and F symmetrical in μ and ρ. (The whole tensor is antisymmetrical in ν and σ.) It is easy to show that

$$F_{\mu\nu\sigma\rho} = (K_{\mu\rho,\nu})_\sigma - (K_{\mu\rho,\sigma})_\nu. \tag{5·9}$$

These, however, are not in-tensors, since the $g_{\mu\nu}$ must be introduced in order to lower the suffix ρ.

Physical Theory.

6. We now introduce the natural gauge of the world, which is determined by measures of space and time made with material and optical appliances. Any apparatus used to measure the world is itself part of the world, so that a natural gauge represents the world as self-gauging. *This can only mean that the tensor $g_{\mu\nu}$ introduced in* (4·1) *is not extraneous and arbitrary, but is a tensor already contained in the world-geometry.* There is only one such tensor available, viz., $^*G_{\mu\nu}$. Hence natural length is given by

$$l^2 = {}^*G_{\mu\nu} A^\mu A^\nu. \tag{6·1}$$

Clearly the antisymmetrical part drops out in this equation, so that

$$l^2 = R_{\mu\nu} A^\mu A^\nu.$$

Accordingly we must take

$$\lambda g_{\mu\nu} = R_{\mu\nu}, \tag{6·2}$$

introducing a universal constant, λ, for convenience, in order to remain free to use the centimetre instead of the natural unit of length.

The difference between $R_{\mu\nu}$ and $G_{\mu\nu}$ in (5·2) consists of terms formed out of the tensor $K_{\mu\nu}{}^\sigma$. We shall find later that this tensor represents electromagnetic and electronic effects, so that the more "empty" the space the more nearly does $R_{\mu\nu}$ approximate to $G_{\mu\nu}$. Hence, in an empty region, the gauging-equation (6·2) becomes

$$G_{\mu\nu} = \lambda g_{\mu\nu}. \tag{6·4}$$

This is De Sitter's equation for a spherical (or, rather, a quasi-spherical) world. The gauging-equation is, in fact, an *alias* of the law of gravitation. If λ approaches zero so that the dimensions of the world become infinitely great compared with our unit of measure, we have Einstein's original gravitational equations $G_{\mu\nu} = 0$. In later work he has used the more general form (6·4), giving a naturally curved world ('Report,' § 50).[†]

[†] It may be noticed that $F_{\nu\sigma} = \frac{1}{2}{}^*B_{\rho\nu\sigma}{}^\rho$, a result easily obtained from (5·1), remembering that $B_{\rho\nu\sigma}{}^\rho = 0$. It follows fairly easily that the proportionate change of volume of a four-dimensional element carried by parallel displacement round a small circuit is $F_{\nu\sigma}\,dS^{\nu\sigma}$.

The form of the world given by these equations is only quasi-spherical, because it requires twenty conditions to determine a truly spherical world, just as it requires twenty conditions ($B_{\mu\nu\sigma\rho} = 0$) to determine a truly flat world. But a world conditioned by (6·4) will have the more important spherical properties. It has this symmetrical form, not from any innate tendency, but because our natural gauge of length at different points and in different directions is adapted to make it so. If we have two electrons at different places, we expect them to be of the same size; we also expect an electron to be spherical. But the equilibrium shape and size of an electron must be determined by something in the world-conditions where it is (apart possibly from small accidental variations persisting from its previous history). The electron must measure itself against something in order to find out how large it ought to be; and in empty space there seems to be nothing but the radii of curvature of the world for it to measure itself against. Similarly, any material object of specified constitution must have determined size and shape *in relation to the radii of the world.* And in so far as the measure of length rests on such bodies, the form of the world expressed in this measure is bound to be quasi-spherical.

It is only in empty space that the approximate equation (6·4) is valid. Electrons are recognised to be small regions of intense local curvature; this is provided for in the accurate gauging equation (6·2), since the difference between $R_{\mu\nu}$ and $G_{\mu\nu}$ may in certain circumstances become large. Intense curvature (or mass) must always be associated with large values of $S_{\mu\nu}{}^{\sigma}$, and we shall find later that large electric fields are associated with large values of this tensor.

Natural space and time must obey Riemannian geometry. It would be easy to devise fanciful geometries in which, for example, ds^4 was a quartic function of the coordinate-differences. But there is only one in-invariant of a simple character[‡] associated with a displacement A^{μ}, viz., $*G_{\mu\nu}A^{\mu}A^{\nu}$, so that (except for far-fetched combinations) no quartic function of the coordinate-differences could have the necessary property. Even $*B_{\mu\nu\sigma}{}^{\rho}$ is no use, for we cannot counteract the upper ρ.

7. Having thus identified physical space-time, we proceed to identify "things." There are three sorts of physical attributes of things to be identified; and our plan is first to inquire what experimental properties

---

[†] The difference between Einstein's and De Sitter's theories depends on the total amount of matter in the world. Both adopt (6·4) as the condition for empty space.

[‡] The next simplest in-invariant is $*B_{\mu\nu\sigma}{}^{\rho}\,*B_{\lambda\tau\rho}{}^{\sigma}\,A^{\mu}A^{\nu}A^{\lambda}A^{\tau}$.

they possess, and then to seek a geometrical tensor possessing these properties *by virtue of mathematical identities.* The tensors need not now be in-tensors, because our description of the experimental properties presupposes the use of the natural gauge-system; and there is no reason to think it would hold good for other gauge-systems:—

(1) The energy-tensor T_μ^ν comprises the energy, momentum, and stress in a region. These must satisfy the condition of conservation of energy and momentum, which is expressed by

$$T_{\mu\ \nu}^{\ \nu} = 0. \tag{7·1}$$

This enables us to make the identification ('Report,' § 39)

$$-8\pi T_\mu^\nu = G_\mu^\nu - \tfrac{1}{2} g_\mu^\nu (G-2\lambda), \tag{7·2}$$

which satisfies (7·1) identically. Here λ might be any constant so far as conservation is concerned. We consider, however, that the energy-tensor should be something which vanishes altogether in empty space (containing no electromagnetic fields). If $T_\mu^\nu = 0$, we have, by contracting (7·2),

$$G = 4\lambda$$

so that

$$0 = G_\mu^\nu - g_\mu^\nu \lambda,$$

or lowering the suffix

$$G_{\mu\nu} = \lambda g_{\mu\nu},$$

agreeing with (6·4) if λ is the same constant. Thus, from the conservation of energy and momentum, we deduce the law of gravitation once more.

(2) The electromagnetic force-tensor $F_{\mu\nu}$ must satisfy two of Maxwell's equations

$$\frac{\partial F_{\mu\nu}}{\partial x_\sigma} + \frac{\partial F_{\nu\sigma}}{\partial x_\mu} + \frac{\partial F_{\sigma\mu}}{\partial x_\nu} = 0. \tag{7·3}$$

This will be an identity if $F_{\mu\nu}$ is the curl of a vector. This leads naturally to an identification with the in-tensor already called $F_{\mu\nu}$, which we have seen is the curl of a vector κ_μ. Hence κ_μ will be the electromagnetic potential.

(3) The electric charge-and-current vector J^μ must satisfy the condition of conservation of electric charge, viz., its divergence $J^\mu_{\ \mu}$ vanishes. This will be an identity if J^μ is itself the divergence of any antisymmetrical contravariant tensor. Accordingly, we identify

$$J^\mu = F^{\mu\nu}_{\ \ \ \nu}. \tag{7·4}$$

Additional confirmation is given by the fact that (7·4) represents Maxwell's two remaining equations ('Report,' § 45).

It would in any case be reasonable to make these identifications—the simplest possible—checking them by comparison with observation later on. But actually very little choice is possible; that is to say, it is not easy to find any other geometrical tensors having the necessary properties as identities. Further, we can scarcely expect much further experimental test, since (with one exception†) all the known laws of mechanics and electromagnetism are already satisfied.

As regards alternatives, some other tensor functions of $g_{\mu\nu}$ are known which satisfy (7·1) identically: but these do not agree with the condition for empty space, $G_{\mu\nu} = \lambda g_{\mu\nu}$. They are extremely complicated expressions, involving derivatives of the fourth or higher orders. An alternative identification of the electromagnetic force with the curl of λ_μ (instead of the curl of κ_μ) seems possible, though rather perverse—seeing that it would leave the fundamental in-tensor $F_{\mu\nu}$ apparently doing nothing to justify its existence.

8. The gauging-equation (6·2) can also be demonstrated from optical theory. The light pulse diverging from a point occupies a conical locus in four dimensions, and must therefore satisfy an equation of the form $a_{\mu\nu}dx_\mu dx_\nu = 0$. This unique locus exists independently of gauge or coordinates, and therefore there must be an in-invariant equation defining it. Clearly $a_{\mu\nu}$ must be an in-tensor, which leaves as the only possibility $a_{\mu\nu} = {}^*G_{\mu\nu}$, or

$$R_{\mu\nu}dx_\mu dx_\nu = 0. \tag{8·1}$$

But in Einstein's theory the light-cone is

$$g_{\mu\nu}dx_\mu dx_\nu = 0, \tag{8·2}$$

so that Einstein's gauge is given by

$$R_{\mu\nu} = \lambda g_{\mu\nu}.$$

Our gauging-equation is therefore certainly true wherever light is propagated, *i.e.*, everywhere except inside the electron. Who shall say what is the ordinary gauge inside the electron? It should, however, be noticed that this optical treatment does not prove that λ is a universal constant. If it were a varying function of position, (8·1) and (8·2) would still agree. The constancy of λ depends on the arguments previously given.

If a displacement A^μ is taken by parallel displacement round a small complete circuit, its change of length is

$$\begin{aligned} \delta l^2 &= \delta\,(g_{\mu\rho}A^\mu A^\rho) = \delta\,(A_\mu A^\mu) = 2\,A_\mu\delta A^\mu \\ &= A_\mu{}^*B_{\rho\nu\sigma}{}^\mu A^\rho dS^{\nu\sigma} \text{ by (3·4)} \\ &= (R_{\rho\nu\sigma\mu} + F_{\rho\nu\sigma\mu})\,A^\mu A^\rho dS^{\nu\sigma} \text{ by (5·8)} \\ &= F_{\rho\nu\sigma\mu}A^\mu A^\rho dS^{\nu\sigma}, \end{aligned} \tag{8·3}$$

since $R_{\rho\nu\sigma\mu}$ is antisymmetrical in μ and ρ.

---

† The law of ponderomotive force of an electromagnetic field, see § 9.

In Weyl's theory $F_{\rho\nu\sigma\mu}$ is of the form $g_{\rho\mu}F_{\nu\sigma}$ so that (8·3) reduces to

$$\delta l^2 = F_{\nu\sigma}dS^{\nu\sigma} \, . \, l^2, \tag{8·4}$$

so that, although a vector in general alters length on describing a circuit, zero-length is unaltered on Weyl's theory. Accordingly, for him zero-length is unique and involves no specification of a route of comparison; whereas in my theory zero-length may not remain zero on parallel displacement according to (8·3). This was for a long while my chief obstacle in generalising his theory, since the abandonment of the uniqueness of zero-length might well seem to leave no unique track for the propagation of light. But we see that the difficulty is surmounted; the unique light-cone is given by the only simple in-invariant equation that can be formed; and the condition that it corresponds to zero-length is one of the factors which determine our natural gauge-system.

9. The ponderomotive force of an electromagnetic field on matter containing electric charges is

$$k_\mu = F_{\mu\nu}J^\nu, \tag{9·1}$$

('Report,' equation (46·1)). This is the tensor-generalisation of the elementary law

$$m\frac{d^2x}{dt^2} = eX, \tag{9·2}$$

which follows immediately from the usual definition of electric force. We have to show that (9·1) can also be deduced when the electric force is defined by the method of identification in § 7.

It will be sufficient to consider an isolated electron. If we consider the experimental verification of (9·1) or (9·2) by the deflection of an electron travelling in an electromagnetic field, it is clear that the field $F_{\mu\nu}$ referred to in (9·1) is the applied external field, and no account is taken of any disturbance of this field caused by the accelerated electron itself. We denote this external field by $F'_{\mu\nu}$. Also, for a particle,

$$\begin{aligned} k^\mu &= -m\left\{\frac{d^2x_\mu}{ds^2} + \{\alpha\beta,\mu\}\frac{dx_\alpha}{ds}\frac{dx_\beta}{ds}\right\} \\ &= -mA^\nu A^\mu{}_\nu, \end{aligned} \tag{9·3}$$

where $A^\mu = dx_\mu/ds$ ('Report,' § 29). If in (9·3) we put $k_\mu = 0$, we obtain the usual equations for the track of an undisturbed particle, *i.e.*, a geodesic, which can be deduced directly from the conservation of energy and momentum.

The equation to be proved is thus

$$mA^\nu A_{\mu\nu} = -F_{\mu\nu}'J^\nu. \tag{9·4}$$

In the ordinary text-books it is proved that the potentials due to a slowly-moving elementary charge de are

$$\mathrm{F}, \mathrm{G}, \mathrm{H}, \Phi = \frac{de}{r}\left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}, 1\right).$$

These expressions are deduced solely from Maxwell's equations (7·3) and (7·4), and we need not here trouble about retarded values, since our application is to the case when r is very small. The general vector-equations are evidently

$$\kappa^\mu = \frac{de}{r}\,\mathrm{A}^\mu, \tag{9·5}$$

where de/r is treated as an invariant, that is to say, its value is to be calculated for Galilean axes with respect to which the electron is momentarily at rest. Hence the electric force *due to the electron* is

$$\mathrm{F}_{\mu\nu} = \kappa_{\mu\nu} - \kappa_{\nu\mu} = (\mathrm{A}_{\mu\nu} - \mathrm{A}_{\nu\mu})\int\frac{de}{r},$$

since all parts of the electron have the same velocity A^μ.

Now, suppose that the electron moves in such a way that its own field on the average just neutralises the external field $\mathrm{F}_{\mu\nu}'$ in the region occupied by the electron. The average value of $\mathrm{F}_{\mu\nu}$ for all the elements of charge constituting the electron will be given by

$$\begin{aligned} e \times \text{average } \mathrm{F}_{\mu\nu} &= (\mathrm{A}_{\mu\nu} - \mathrm{A}_{\nu\mu})\iint\frac{de_1 de_2}{r_{12}} \\ &= (\mathrm{A}_{\mu\nu} - \mathrm{A}_{\nu\mu})\, e^2/a, \end{aligned}$$

where $1/a$ is an average value of $1/r$ taken for every pair of points in the electron. Also $\mathrm{J}^\mu = e\mathrm{A}^\mu$. Hence, multiplying by A^ν, we shall have, if the external field is neutralised,

$$\mathrm{F}_{\mu\nu}'\mathrm{J}^\nu = -\mathrm{A}^\nu(\mathrm{A}_{\mu\nu} - \mathrm{A}_{\nu\mu})\, e^2/a.$$

But $$\mathrm{A}^\nu\mathrm{A}_{\nu\mu} = \mathrm{A}^\nu{}_\mu\mathrm{A}_\nu = \tfrac{1}{2}(\mathrm{A}^\nu\mathrm{A}_\nu)_\mu = 0,$$

since $$\mathrm{A}^\nu\mathrm{A}_\nu = g_{\mu\nu}\frac{dx_\mu}{ds}\frac{dx_\nu}{ds} = 1.$$

Hence $$\mathrm{F}_{\mu\nu}'\mathrm{J}^\nu = -\mathrm{A}^\nu\mathrm{A}_{\mu\nu}\, e^2/a, \tag{9·6}$$

which agrees with (9·4) if the mass of the electron is

$$m = e^2/a. \tag{9·7}$$

It is now evident that (9·1) implies that an electron is a structure which cannot exist in a resultant field of electromagnetic force. Further than this we cannot expect to go without a detailed analysis of electron structure. It is a question of the adjustment of the electron to the world-conditions in which it finds itself. The electron evidently has a certain amount of adaptability to different world-conditions; and can always resist differential electromagnetic force, just as a static electron can resist the radial electrostatic force due to its own charge. Knowing nothing of the laws of electron structure, the existence of an isolated electron at rest in no external field is apparently a "miracle"; our calculation shows that the existence of an electron having the acceleration given by (9·1) in an external electromagnetic field is *precisely the same miracle* viewed from another standpoint.

There is a close resemblance between this method of deducing the acceleration of an electron in an electric field and that of a material body in a gravitational field by the principle of equivalence. Neither body can be in equilibrium (*i.e.*, exist) if there is any resultant field of force relative to it; it would be out of adjustment with the surrounding state of the world, and the hypothesis represents a self-contradictory structure. Therefore the state of motion must be such that no resultant field of force exists where the object is; the body and the electron destroy the field by their own accelerations, and this criterion gives the motion in each case. Both, however, are able in some way to oppose differential forces by internal stresses.

The electric mass m is not immediately a source of gravitational attraction, so far as the foregoing discussion can show. But an electron *has* gravitational mass; it is by hypothesis a place where $F_{\mu\nu}$, and therefore $S_{\mu\nu}{}^\sigma$, has very large components. Reference to (5·4) shows that in this case G will differ very considerably from R (unless there is an unsuspected cancelling of terms), showing that the electron is a region of abnormal world-curvature, and therefore the source of gravitational effects. From the constancy of structure of electrons, the gravitational mass of a pair of electrons (positive and negative) will be in a constant ratio to their electric mass. We are under no obligation to prove that the ratio is the same for the positive and negative electrons separately; because there is no observational evidence for this. In §11, however, it appears likely that the ratio of the electrical to the gravitational mass is $8\pi\lambda$; if so, it may be presumed that the factor applies to positive and negative electrons independently.

From (9·1) the expression for the electromagnetic energy-tensor is deduced as explained in 'Report,' §46.

Further Developments.

10. The whole energy-tensor is given by

$$-8\pi T_\mu{}^\nu = \lambda\left(G_\mu{}^\nu - \tfrac{1}{2} g_\mu{}^\nu (G - 2\lambda)\right), \tag{10·1}$$

and the electromagnetic energy-tensor by

$$E_\mu{}^\nu = -F_{\mu\sigma}F^{\nu\sigma} + \tfrac{1}{4} g_\mu{}^\nu F_{\alpha\beta}F^{\alpha\beta}, \tag{10·2}$$

('Report,' equations (50·4) and (46·4)). It is necessary to insert the factor λ in (10·1) in order to make the dimensions consistent. Even now it is not certain that the two expressions are in common units; but, at any rate, their ratio will not be altered by changing the unit of length, *i.e.*, changing λ.

Contracting (10·2), $E = 0$ identically. It is thus impossible to build up matter or electrons from electromagnetic fields alone, and some other form of energy must be present. This represents the non-Maxwellian binding forces. It is convenient to write

$$R_{\mu\nu} = G_{\mu\nu} + H_{\mu\nu}, \qquad R = G + H, \tag{10·3}$$

so that from (5·2) and (5·4)

$$H_{\mu\nu} = \kappa_{\mu\nu} + \kappa_{\nu\mu} - (S_{\mu\nu}{}^{\sigma})_{\sigma} + \{S_{\alpha\nu}{}^{\beta} S_{\beta\mu}{}^{\alpha} - 2\kappa_{\alpha} S_{\mu\nu}{}^{\alpha}\}, \tag{10·4}$$

$$H = 2\kappa^{\alpha}{}_{\alpha} + \lambda^{\alpha}{}_{\alpha} + \{S_{\alpha\beta,\gamma} S^{\alpha\gamma,\beta} + 2\kappa_{\alpha}\lambda^{\alpha}\}. \tag{10·5}$$

Contracting (10·1)

$$-8\pi T = \lambda(-G + 4\lambda).$$

But

$$R = g^{\mu\nu} R_{\mu\nu} = g^{\mu\nu} \,.\, \lambda g_{\mu\nu} = 4\lambda.$$

Hence

$$-8\pi T = \lambda(R - G) = \lambda H, \tag{10·6}$$

and similarly

$$8\pi T_{\mu}{}^{\nu} = \lambda(H_{\mu}{}^{\nu} - \tfrac{1}{2} g_{\mu}{}^{\nu} H). \tag{10·7}$$

We may regard these formulæ as giving the electrical aspect of material energy relating it to the tensor $S_{\mu\nu}{}^{\sigma}$, as opposed to (10·1), which shows its gravitational aspect.

To fix ideas as to the order of magnitude of terms, let us take the centimetre as unit of length, and choose approximately Galilean coordinates with a centimetre mesh. In these units λ (the curvature of the world in empty space) can scarcely be greater than 10^{-50}. Calling this a small quantity of the first order, it is easily seen that $S_{\mu\nu}{}^{\sigma}$ is of the first order, except in electric fields so intense as to alter the order of magnitude of the local curvature. Accordingly, the expressions for $H_{\mu\nu}$ and H break up into first-order and second-order terms, the latter being bracketed in (10·4) and (10·5). Evidently the formation of electrons of definite size is associated with the non-linearity of these equations, the product terms becoming important when the electric field is sufficiently intense.

For an electron at rest $\kappa^{\alpha}{}_{\alpha}$ vanishes. Although the physical interpretation of λ^{α} is not known, it seems likely that $\lambda^{\alpha}{}_{\alpha}$ will also vanish in the static case. Thus the mass of an isolated electron is probably represented entirely by the product terms in H.

11. It is natural to infer from (10·1) and (10·2) that in space containing no electrons $T_{\mu}{}^{\nu}$ and $E_{\mu}{}^{\nu}$, if reduced to a common unit, must be equal. But this involves an arbitrary localisation of the energy-tensors which is scarcely

justified, since the connection established between them *viâ* (9·1) refers to the integrated expressions.

The difficulty of localisation can be realised by considering an electron at rest. In this case its relative mass is equal to its invariant mass. On the ordinary view the relative mass is attributed to the inertia of the electromagnetic field which surrounds it; but the field contributes nothing to the invariant mass. The two masses are equal, if not synonymous, yet one is arbitrarily located in the field and the other in the electron.

From the formulæ already given it is possible to express T_μ^ν and E_μ^ν in terms of the S or K tensor; but no obvious connection between them can be found. It is possible that the formulæ may admit of some simplification, since the gauging equation must introduce some relations between the components of $K_{\mu\nu}^\sigma$, which, however, are difficult to follow out. The following discussion seems to arrive at a natural connection between T_μ^ν and E_μ^ν, and it appears to have a fundamental significance.

Consider the invariant

$$U = \int {}^*G_{\mu\nu}{}^*G^{\nu\mu}\sqrt{(-g)}\,d\tau, \tag{11·1}$$

which, though not strictly an in-invariant, is independent of the value of λ, and is accordingly a pure number.[†] We have

$$\begin{aligned} {}^*G_{\mu\nu}{}^*G^{\nu\mu} &= (R_{\mu\nu}+F_{\mu\nu})(R^{\nu\mu}+F^{\nu\mu}) \\ &= R_{\mu\nu}R^{\mu\nu}-F_{\mu\nu}F^{\mu\nu}, \end{aligned} \tag{11·2}$$

owing to the symmetry of $R_{\mu\nu}$ and antisymmetry of $F_{\mu\nu}$.

Calculate the variation of U due to small arbitrary variations of the $g_{\mu\nu}$ which vanish at and near the boundary of the region considered. (These variations represent alterations of the state of the world expressed by the in-tensor $R_{\mu\nu}$, which lead to corresponding variations of the $g_{\mu\nu}$ through the gauging-equation.) We have

$$\begin{aligned} \delta[R_{\mu\nu}R^{\mu\nu}\sqrt{(-g)}] &= R_{\mu\nu}\delta[R^{\mu\nu}\sqrt{(-g)}] \\ &\qquad + R^{\mu\nu}\delta[R_{\mu\nu}\sqrt{(-g)}]-R_{\mu\nu}R^{\mu\nu}\delta[\sqrt{(-g)}] \\ &= \lambda\{g_{\mu\nu}\delta[R^{\mu\nu}\sqrt{(-g)}]+g^{\mu\nu}\delta[R_{\mu\nu}\sqrt{(-g)}]\} \\ &\qquad -4\lambda^2\delta[\sqrt{(-g)}], \end{aligned}$$

but $\lambda\sqrt{(-g)}(R^{\mu\nu}\delta g_{\mu\nu}+R_{\mu\nu}\delta g^{\mu\nu}) = \lambda^2\sqrt{(-g)}(g^{\mu\nu}\delta g_{\mu\nu}+g_{\mu\nu}\delta g^{\mu\nu}) = 0$,

so that, by adding,

$$\begin{aligned} \delta[R_{\mu\nu}R^{\mu\nu}\sqrt{(-g)}] &= 2\lambda\delta[R\sqrt{(-g)}]-4\lambda^2\delta[\sqrt{(-g)}] \\ &= 2\lambda\delta\{(G-2\lambda)\sqrt{(-g)}\}+2\lambda\delta[H\sqrt{(-g)}] \text{ by (10·3).} \end{aligned}$$

[†] As Weyl has pointed out, the existence of several numerical invariants of a simple character is a distinctive property of a *four*-dimensional continuum. The invariant U is from many points of view the most fundamental of these.

But $(G-2\lambda)\sqrt{(-g)}$ is well known as the "action-density" of the gravitational field, and its variation is reducible to[†]

$$\sqrt{(-g)}\,\{G_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}(G-2\lambda)\}\,\delta g^{\mu\nu} = -8\pi T_{\mu\nu}\sqrt{(-g)}\,\delta g^{\mu\nu}/\lambda,$$

rejecting a part which yields a surface-integral over the boundary, since all the variations vanish there. Again,

$$\delta[F_{\mu\nu}F^{\mu\nu}\sqrt{(-g)}] = F_{\mu\nu}F_{\alpha\beta}\,\delta[g^{\alpha\mu}g^{\beta\nu}\sqrt{(-g)}]$$

since the in-tensor $F_{\mu\nu}$ is not affected by the variations here considered. It is readily found[‡] that this reduces to $-2E_{\mu\nu}\sqrt{(-g)}\,\delta g^{\mu\nu}$.

Further by (10·5)

$$\delta\int H\sqrt{(-g)}\,d\tau = \delta\int\{S_{\alpha\beta,\gamma}S^{\alpha\gamma,\beta}+2\kappa_\alpha\lambda^\alpha\}\sqrt{(-g)}\,d\tau.$$

The first-order terms are omitted, since

$$(2\kappa^\alpha{}_\alpha+\lambda^\alpha{}_\alpha)\sqrt{(-g)} = \frac{\partial}{\partial x_\alpha}\{2\kappa^\alpha\sqrt{(-g)}+\lambda^\alpha\sqrt{(-g)}\},$$

by a well-known formula for the divergence of a contravariant vector. This on integration gives a surface-integral over the boundary where all variations vanish.

Collecting these results we have

$$-\tfrac{1}{2}\delta U = \int(8\pi T_{\mu\nu}-E_{\mu\nu})\,\delta g^{\mu\nu}\,.\,\sqrt{(-g)}\,d\tau-\lambda\,\delta\int\{S_{\alpha\beta,\gamma}S^{\alpha\gamma,\beta}+2\kappa_\alpha\lambda^\alpha\}\sqrt{(-g)}\,.\,d\tau. \tag{11·3}$$

Thus, by considering the variation of what is probably the most fundamental numerical invariant, we have brought together in one formula the whole energy-tensor and the electro-magnetic energy-tensor—a result which would be difficult to accomplish in any other way. There can be little doubt that in the combination

$$8\pi T_{\mu\nu}-E_{\mu\nu}, \tag{11·4}$$

they are expressed in a common unit, so that the difference must represent the electronic or non-Maxwellian part of the whole energy-tensor. The remaining term on the right is also concerned with electronic terms, since outside the electron the integrand is of the second order of small quantities. Hence δU must be a variation of "action" associated with electrons. It is not clear whether we are here approaching a law of Nature or whether our analysis is a device for bringing out hidden identities in the mathematical expressions; but I incline to the latter view.

[†] 'Raum, Zeit, Materie,' 3rd edition, pp. 205, 253.

[‡] *Ibid.*, p. 197.

It may be tempting to suggest a law of stationary action that $\delta U = 0$ This, however, is impossible. If we do not split up R into $G+H$, we find

$$-\tfrac{1}{2}\delta U = \int (\lambda^2 g_{\mu\nu} - E_{\mu\nu})\,\delta g^{\mu\nu}\,.\,\sqrt{(-g)}\,d\tau,$$

and our proposed law of action would be inconsistent with the identity $E = 0$.

It need not surprise us that the combination (11·4) is only met with when the reduction proceeds in a particular way. As I have already indicated, the real connection is between the integrated energies, and the localisation is probably not such as to give the same connection in each element of volume. I think that the casting away of the surface-integral from $\delta[G\sqrt{(-g)}]$ has redistributed the location so as to display the agreement.

The most important result is the reduction of $T_{\mu\nu}$ and $E_{\mu\nu}$ to a common unit, implied in (11·4). $T_{\mu\nu}$ for given masses or electric fields is known in ordinary units, whereas $E_{\mu\nu}$ involves the natural units which determine the in-tensor $F_{\mu\nu}$. Thus, taking $\lambda = 10^{-50}$ as an upper limit, I find that the quantum is 10^{-114} natural units of action. A faint hope that the natural unit of action involved in this theory might prove to be identical with the quantum is thus disappointed. It appears also that $R_{\mu\nu}$ and $F_{\mu\nu}$ may be of roughly equal magnitude in electric fields of quite ordinary strength produced in the laboratory.

12. It is a natural impression that in the light of the more general theory Einstein's work must be regarded as only a close approximation. The impression is wrong, for the present paper leads to the conclusion that Einstein's postulates and deductions are exact. The natural geometry of the world (in the sense that I have defined in 'Space, Time, and Gravitation') is the geometry of Riemann and Einstein, not Weyl's generalised geometry or mine. What *we* have sought is not the geometry of actual space and time, but the geometry of the world-structure, which is the common basis of space and time and things. I have tried to emphasise this distinction by dividing the paper into Geometrical and Physical theory. The first part is the so-called "geometry" of a general kind of relation complex; for I think that any conception of *structure* (as opposed to substance) must be analysable into a complex of relations and relata, the relata having no structural significance except as the meeting point of several relations, and the relations having no significance except as connecting and ordering the relata. This part of the theory is an essay in assigning measure to such a structure. Our method is based on an axiom of parallel displacement, *i.e.*, comparability of proximate relations; this seems to be the minimum degree of comparability which permits of any differentiation of structure. We succeed in arriving at a quantitative measure of this structure, viz., $^*G_{\mu\nu}$,

which breaks up into symmetrical and antisymmetrical parts, $g_{\mu\nu}$ and $F_{\mu\nu}$, the latter being necessarily the curl of a vector. It is at this bifurcation of the theory that Einstein begins; he has recognised these two tensors from their physical manifestations, and he expresses no view as to what lies behind them. From $g_{\mu\nu}$ he builds up interval-length, and hence actual space, time, geometry, gravitation, and mechanics; from $F_{\mu\nu}$ he builds the Maxwellian theory of electricity and magnetism. His *interval* is absolute, for it is our in-invariant $R_{\mu\nu}dx_\mu dx_\nu$; so that *his work stands unaffected by the ambiguity of length-comparisons in the initial stages of our theory.* The material he employs is furnished by our theory in exact accordance with his specification.

We have further shown that there is a more comprehensive measure of the structure given by $^*B_{\mu\nu\sigma}{}^{\rho}$, of which $^*G_{\mu\nu}$ is an abbreviated summary. It seems obvious that even this extended tensor is by no means an exhaustive description of all that there is in the world; but presumably it (or something analogous) marks the limit of physical science, though not necessarily of the scientific method. The curious point is that, for present-day physics, the summary $^*G_{\mu\nu}$ suffices, and either the extended $^*B_{\mu\nu\sigma}{}^{\rho}$ has no message for us or we have failed to interpret the message. Even in § 11 the invariant U belongs to the summary description. It remains to be seen whether there is anything of observational importance omitted in the summary—anything which plays a part in the formation of negative and positive electrons, and possibly even of the quantum.

The Field of an Electron on Einstein's Theory of Gravitation.

By G. B. JEFFERY, M.A., D.Sc., Fellow of University College, London.

(Communicated by Prof. L. N. G. Filon, F.R.S. Received December 16, 1920,—Revised January 11, 1921.)

Einstein's Generalised Theory of Relativity has accomplished notable results in the region of astronomical mechanics, but for the moment it seems difficult to go further in this direction until some further progress is made with the theory of the solution of the field equations. Practically all the consequences of the theory which have been established so far are obtained from the solution of the equations corresponding to a single isolated singularity. The *exact* solutions corresponding to two isolated singularities are urgently required for the further exploration of the theory. The approximate solutions which have been put forward by De Sitter,* Dröste† and Einstein,‡ though valuable in default of exact solutions, are apt to be misleading and perhaps to exclude effects of far-reaching theoretical importance such as radiation. Meanwhile it is well to examine the consequences of the theory at the other extreme of the realm of physical science, in its relation to atomic phenomena. Here it seems no longer justifiable to consider gravitation and electricity separately. In these microscopic phenomena, where we are free from the effects of averaging, mass and charge seem to be inextricably connected. They exist in certain definite combinations, possibly in only two combinations as electrons and hydrogen nuclei. The gravitational field corresponding to such a charged particle has been investigated by Nordström§ by the application of the calculus of variations to the Hamiltonian function of the combined fields. Nordström's work was not brought to our notice until after the present paper was written and we had obtained the same result by direct solution of the field equations. As the methods employed may perhaps be more familiar to English students of the subject it has been thought well to give a brief outline of this proof as an alternative to Nordström's in the first section of this paper.

---

* 'Monthly Notices of the Royal Astronomical Society,' vol. 72, p. 155 (1916).

† 'Proc. Acad. Amsterdam,' vol. 19, p. 447 (1916).

‡ 'Berlin, Sitzungsberichte,' 1916, p. 688.

§ "On the Energy of the Gravitational Field in Einstein's Theory," 'Proc. Ac. Amsterdam,' vol. 20, p. 1236 (1918).

§ 1. *The Solution of the Field Equations.*

The gravitational field is defined by the symmetrical covariant tensor $g_{\mu\nu}$, where

$$ds^2 = \Sigma\Sigma(\mu, \nu) g_{\mu\nu}\, dx_\mu\, dx_\nu \qquad (\mu, \nu = 1, 2, 3, 4).$$

where the variables of summation are enclosed in brackets immediately following the signs of summation.

For a symmetrical static field we may, following Eddington* and Schwartzchild, take polar co-ordinates r, θ, ϕ, $\tau = ct$ and

$$ds^2 = -e^\lambda\, dr^2 - r^2\, d\theta^2 - r^2 \sin^2\theta\, d\phi^2 + e^\nu\, d\tau^2, \tag{1}$$

where λ, ν are functions of r only, so that $g_{\mu\nu} = 0$ if $\mu \neq \nu$ and

$$g_{11} = -e^\lambda, \quad g_{22} = -r^2, \quad g_{33} = -r^2 \sin^2\theta, \quad g_{44} = e^\nu, \tag{2}$$

The determinant of $g_{\mu\nu}$ is given by

$$g = -r^4 \sin^2\theta\, e^{\lambda+\nu}. \tag{3}$$

and the associated contravariant tensor by

$$g^{\mu\mu} = 1/g_{\mu\mu} \qquad \text{and} \qquad g^{\mu\nu} = 0 \text{ for } \mu \neq \nu. \tag{4}$$

The Christoffel symbols, defined by

$$\{\lambda\mu, \nu\} = \tfrac{1}{2}\Sigma(\alpha) g^{\nu\alpha}\left(\frac{\partial g_{\lambda\alpha}}{\partial x_\mu} + \frac{\partial g_{\mu\alpha}}{\partial x_\lambda} - \frac{\partial g_{\lambda\mu}}{dx_\alpha}\right),$$

are then as given by Eddington,† accents denoting differentiation with respect to r,

$$\left.\begin{array}{lll} \{11, 1\} = \tfrac{1}{2}\lambda', & \{12, 2\} = \{13, 3\} = 1/r, & \{14, 4\} = \tfrac{1}{2}\nu', \\ \{22, 1\} = -re^{-\lambda}, & \{23, 3\} = \cot\theta, & \\ \{33, 1\} = -r\sin^2\theta e^{-\lambda}, & \{33, 2\} = -\sin\theta\cos\theta, & \\ & \{44, 1\} = \tfrac{1}{2}\nu' e^{\nu-\lambda}, & \end{array}\right\}, \tag{5}$$

the rest vanishing identically.

The components of Einstein's contracted tensor $G_{\mu\nu}$ vanish with the exception of

$$\left.\begin{array}{l} G_{11} = \tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 - \lambda'/r. \\ G_{22} = e^{-\lambda}\{1 + \tfrac{1}{2}r(\nu' - \lambda')\} - 1. \\ G_{33} = \sin^2\theta\,[e^{-\lambda}\{1 + \tfrac{1}{2}r(\nu' - \lambda')\} - 1] \\ G_{44} = -e^{\nu-\lambda}\{\tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 + \nu'/r\}. \end{array}\right\}. \tag{6}$$

---

* 'Report on Relativity to the Physical Society,' second edition, p. 44. Eddington's Methods have been closely followed in this paper, and references to this source will be quoted as 'Report.'

† 'Report,' p. 45.

The electromagnetic field is defined* by the covariant vector

$$\kappa_\mu \equiv (-F, -G, -H, \Phi)$$

where F, G, H, are the components of the vector potential and Φ is the scalar potential. For static symmetry about the origin $n_2 = n_3 = 0$, and n_1, n_4 are functions of r only.

Denoting the covariant derivative of κ_μ by $\kappa_{\mu\sigma}$ the electromagnetic covariant tensor is defined by

$$F_{\mu\sigma} = \kappa_{\mu\sigma} - \kappa_{\sigma\mu},$$

and in our case all the components vanish with the exception of

$$F_{14} = -\kappa_4', \qquad F_{41} = \kappa_4', \tag{7}$$

the accents as before denoting differentiation with respect to r.

The associated mixed and contravariant tensors are defined respectively by

$$F_\mu{}^\sigma = \Sigma(\alpha) g^{\sigma\alpha} F_{\mu\alpha}, \qquad F^{\mu\sigma} = \Sigma\Sigma(\alpha, \beta) g^{\mu\alpha} g^{\sigma\beta} F_{\alpha\beta}, \tag{8}$$

and in our case the only surviving components are

$$F_4{}^1 = g^{11} F_{41} = -e^{-\lambda} \kappa_4', \qquad F_1{}^4 = g^{44} F_{14} = -e^{-\nu} \kappa_4' \tag{9}$$

and

$$F^{14} = g^{11} g^{44} F_{14} = e^{-(\lambda+\nu)} \kappa_4' = -F^{41}. \tag{10}$$

In the absence of charge, Maxwell's equations of the electromagnetic field are expressed† by the vanishing of the contracted covariant derivative of $F^{\mu\sigma}$,

$$F_\sigma{}^{\mu\sigma} \equiv \frac{1}{\sqrt{(-g)}} \Sigma(\sigma) \frac{\partial}{\partial x_\sigma} [\sqrt{(-g)}\, F^{\mu\sigma}]$$

in Lorentz units.

These lead to a single equation in n_4,

$$\frac{\partial}{\partial r} (r^2 \kappa_4' e^{-\frac{1}{2}(\lambda-\nu)}) = 0. \tag{11}$$

The mixed electromagnetic energy tensor is defined by

$$E_\sigma{}^\mu = -\Sigma(\alpha) F_{\sigma\alpha} F^{\mu\alpha} + \tfrac{1}{4} g_\sigma{}^\mu \Sigma\Sigma(\alpha, \beta) F^{\alpha\beta} F_{\alpha\beta}.$$

In our case the only surviving components are

$$\left.\begin{aligned} E_1{}^1 = E_4{}^4 &= \tfrac{1}{2} \kappa_4'^2 e^{-(\lambda+\nu)} \\ E_2{}^2 = E_3{}^3 &= -\tfrac{1}{2} \kappa_4'^2 e^{-(\lambda+\nu)} \end{aligned}\right\}. \tag{12}$$

The associated covariant energy tensor

$$E_{\mu\sigma} = \Sigma(\alpha) g_{\alpha\sigma} E_\mu{}^\alpha$$

vanishes except for the components

$$\left.\begin{aligned} E_{11} &= -\tfrac{1}{2} \kappa_4'^2 e^{-\nu}, & E_{22} &= \tfrac{1}{2} r^2 \kappa_4'^2 e^{-(\lambda-\nu)} \\ E_{33} &= \tfrac{1}{2} r^2 \sin^2\theta\, \kappa_4'^2 e^{-(\lambda+\nu)}, & E_{44} &= \tfrac{1}{2} \kappa_4'^2 e^{-\lambda} \end{aligned}\right\}, \tag{13}$$

---

* 'Report,' p. 76.

† 'Report,' p. 77.

and the associated scalar

$$\mathrm{E} = \Sigma\Sigma\,(\mu\sigma)\, g^{\mu\sigma}\,\mathrm{E}_{\mu\sigma} = \Sigma\,(\alpha,\,\mu)\, g_{\alpha}{}^{\mu}\,\mathrm{E}_{\mu}{}^{\alpha} = \mathrm{E}_1{}^1 + \mathrm{E}_2{}^2 + \mathrm{E}_3{}^3 + \mathrm{E}_4{}^4 = 0. \quad (14)$$

Omitting the terms representing the density and motion of matter, Einstein's equations of the gravitational field are

$$\mathrm{G}_{\mu\sigma} = -8\pi\kappa c^{-4}\,(\mathrm{E}_{\mu\sigma} - \tfrac{1}{2} g_{\mu\sigma}\,\mathrm{E}),$$

where κ is the constant of gravitation and c the velocity of light

These give

$$\begin{aligned}
\mathrm{G}_{11} &= 4\pi\kappa c^{-4}\kappa_4'^{2} e^{-\nu},\\
\mathrm{G}_{22} &= -4\pi\kappa c^{-4} r^2 \kappa_4'^{2} e^{-(\lambda+\nu)},\\
\mathrm{G}_{33} &= -4\pi\kappa c^{-4} r^2 \sin^2\theta\,\kappa_4'^{2} e^{-(\lambda+\nu)},\\
\mathrm{G}_{44} &= -4\pi\kappa c^{-4}\kappa_4'^{2} e^{-\lambda},
\end{aligned}$$

the remaining components of $\mathrm{G}^{\mu\sigma}$ vanishing.

Hence from (6) and (15) we have

$$\tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 - \lambda' r = 4\pi\kappa c^{-4}\kappa_4'^{2} e^{-\nu}, \qquad (16)$$

$$e^{-\lambda}\{1 + \tfrac{1}{2} r(\nu' - \lambda')\} - 1 = -4\pi\kappa c^{-4} r^2 \kappa_4'^{2} e^{-(\lambda+\nu)}, \qquad (17)$$

$$\sin^2\theta\,[e^{-\lambda}\{1 + \tfrac{1}{2} r(\nu' - \lambda')\} - 1] = -4\pi\kappa c^{-4} r^2 \sin^2\theta\,\kappa_4'^{2} e^{-(\lambda+\nu)}, \qquad (18)$$

$$\tfrac{1}{2}\nu'^{1} - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 + \nu'/r = 4\pi\kappa c^{-4}\kappa_4'^{2} e^{-\nu}. \qquad (19)$$

Equations (17) and (18) are clearly identical, while (16) and (19) give $\lambda' + \nu' = 0$ or $\lambda + \nu = const.$ If the gravitational field disappears at a great distance from the origin we may take $\lambda + \nu = 0$.

Equation (11) then gives

$$\kappa_4 = \frac{\epsilon}{4\pi r}, \qquad (20)$$

where ϵ is a constant which will presently be identified with the electric charge in Lorentz units, and a second constant of integration has been put equal to zero without loss of generality.

Equation (17) then gives

$$e^{\nu}(1 + r\nu') = 1 - \frac{\kappa\epsilon^2}{4\pi c^4 r^2},$$

or writing $e^{\nu} = \gamma$,

$$\gamma = 1 - \frac{2\kappa m}{c^2 r} + \frac{\kappa\epsilon^2}{4\pi c^4 r^2}, \qquad (21)$$

where m is a constant of integration which will subsequently be seen to be the mass. It may now be verified that equations (16) and (19) are satisfied identically.

We have therefore as the line element of the space surrounding a point which is a singularity of *both* the electric and gravitational fields

$$ds^2 = -\gamma^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta\, d\phi^2 + \gamma c^2 dt^2, \qquad (22)$$

which is the result given by Nordström.

§2. *The Motion of a Particle in the Field of a Charged Nucleus.*

We will investigate the motion of a particle having both charge and inertia mass in the field which we have examined in the previous section. The mass and charge of the particle will be supposed so small that, while the motion of the particle is determined by the field, the field itself is not appreciably affected by the motion of the particle. This is a serious limitation, but it is the best that we can do in default of the solution corresponding to two singularities. It neglects all effects due to radiation and to the reaction upon the nucleus. If the nucleus is itself an electron, it is difficult to imagine any physical phenomenon to which the analysis can apply, unless indeed we regard the electron as composed of still smaller elements in a disposition which is stable for the particular values of the electronic charge, mass and radius. The analysis will, however, give a good approximation to the motion of an electron in the neighbourhood of an atomic nucleus, for then the mass of the nucleus is at least about a thousand times that of the electron, and, while the charges may be of the same order, it will appear that the terms in $g_{\mu\nu}$ which depend on the charge are in all practical cases small compared with those depending upon the mass.

The "force" acting on a charged particle is given by[*] the covariant vector

$$k_\sigma = \mathrm{F}_{\sigma\mu}\,\mathrm{J}^\mu,$$

where J^μ is the charge-current vector which, for a single particle of charge ϵ' is

$$\epsilon'\left(\frac{dx_1}{ds},\ \frac{dx_2}{ds},\ \frac{dx_3}{ds},\ \frac{dx_4}{ds}\right).$$

The associated contravariant vector is

$$k^\sigma = \Sigma\,(\alpha)\,g^{\alpha\sigma}\,k_\alpha. \tag{23}$$

The contravariant acceleration vector[†] is

$$\mathrm{A}^\sigma = \frac{d^2x_\sigma}{ds^2} + \Sigma\Sigma\,(\alpha,\beta)\,\{\alpha\beta,\sigma\}\,\frac{dx_\alpha}{ds}\,\frac{dx_\beta}{ds}. \tag{24}$$

We will take as the equations of motion of our particle

$$m'\,\mathrm{A}^\sigma = -\frac{1}{c^2}\,k^\sigma, \tag{25}$$

where m' is the mass of the particle.

In rectangular co-ordinates $x, y, z, \tau = ct$ in a non-gravitational field, the Christoffel symbols vanish, and if v is the velocity of the particle

$$\frac{d}{ds} = \frac{1}{\sqrt{(c^2-v^2)}}\,\frac{d}{dt}.$$

---

[*] 'Report,' p. 78.

[†] 'Report,' p. 48.

Remembering that the electric and magnetic vectors (E, B) are given by

$$\mathrm{B} = curl\,(\mathrm{F}, \mathrm{G}, \mathrm{H}), \qquad \mathrm{E} = -grad\Phi - \frac{1}{c}\frac{\partial}{\partial t}(\mathrm{F}, \mathrm{G}, \mathrm{H}),$$

the equations (25) are easily reduced to

$$\frac{d}{dt}\left\{\frac{m'}{\sqrt{(1-v^2/c^2)}}\frac{dx}{dt}\right\} = \epsilon'\left(\mathrm{E}_x + \frac{1}{c}(v_y\mathrm{B}_z - v_z\mathrm{B}_y)\right)$$

and two similar equations, with

$$\frac{d}{dt}\left\{m'c^2\left(\frac{1}{\sqrt{(1-v^2/c^2)}}-1\right)\right\} = \epsilon'(\mathrm{E}_x v_x + \mathrm{E}_y v_y + \mathrm{E}_z v_z).$$

These are the equations of motion on the restricted relativity theory, and reduce to the classical equations for small velocities. Equations (25), therefore, hold in the absence of gravitation, and being of the necessary tensor form they will, by the equivalence hypothesis, be true for any system of co-ordinates and in a gravitational field.

In our co-ordinates

$$k_1 = \frac{\epsilon\epsilon' c}{4\pi r^2}\frac{dt}{ds}, \quad k_2 = k_3 = 0, \quad k_4 = -\frac{\epsilon\epsilon'}{4\pi r^2}\frac{dr}{ds}$$

and hence

$$k^1 = -\frac{\epsilon\epsilon' c e^{\nu}}{4\pi r^2}\frac{dt}{ds}, \quad k^2 = k^3 = 0, \quad k^4 = -\frac{\epsilon\epsilon' e^{-\nu}}{4\pi r^2}\frac{dr}{ds}.$$

Substituting from (5) and (24) in (25) we have

$$\frac{d^2r}{ds^2} - \tfrac{1}{2}\nu'\left(\frac{dr}{ds}\right)^2 - re^{\nu}\left(\frac{d\theta}{ds}\right)^2 - r\sin^2\theta\, e^{\nu}\left(\frac{d\phi}{ds}\right)^2 + \tfrac{1}{2}c^2\nu' e^{2\nu}\left(\frac{dt}{ds}\right)^2 = \frac{\epsilon\epsilon' e^{\nu}}{4\pi c m' r^2}\frac{dt}{ds}, \tag{26}$$

$$\frac{d^2\theta}{ds^2} + \frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds} - \sin\theta\cos\theta\left(\frac{d\phi}{ds}\right)^2 = 0, \tag{27}$$

$$\frac{d^2\phi}{ds^2} + \frac{2}{r}\frac{dr}{ds}\frac{d\phi}{ds} + 2\cot\theta\frac{d\theta}{ds}\frac{d\phi}{ds} = 0, \tag{28}$$

$$\frac{d^2t}{ds^2} + \nu'\frac{dr}{ds}\frac{dt}{ds} = \frac{\epsilon\epsilon' e^{-\nu}}{4\pi c^3 m' r^2}\frac{dr}{ds}. \tag{29}$$

For motion in the equatorial plane we may put $\theta = \frac{1}{2}\pi$, and (27) is then identically satisfied. From (28) and (29) we then have

$$r^2\frac{d\phi}{ds} = h, \qquad \gamma\frac{dt}{ds} = n - \frac{\epsilon\epsilon'}{4\pi c^3 m' r}, \tag{30}$$

where h and n are constants. Eliminating ϵ, ϵ' between (26) and (29) we obtain a third integral

$$\frac{1}{\gamma}\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\phi}{ds}\right)^2 - c^2\gamma\left(\frac{dt}{ds}\right)^2 + 1 = 0, \tag{31}$$

where γ is given by (21).

Considering a particle at rest in three dimensional space so that $\frac{dr}{ds} = \frac{d\phi}{ds} = 0$, we obtain from (26) and (31)

$$m' \frac{d^2 r}{dt^2} = -\kappa \frac{mm'}{r^2} + \frac{\epsilon\epsilon'}{4\pi r^2}$$

where terms of higher order than r^{-2} have been neglected. Hence this is the equation of motion of a particle at great distances from the origin, and our interpretation of the constants m, ϵ as the mass and charge of the nucleus respectively, is confirmed.

We have used Lorentz units in order to maintain symmetry among the components of the various tensors employed. We may now revert to ordinary electrostatic units by omitting the factor 4π in (21) and (30). We then have as our equations of motion,

$$\frac{1}{\gamma}\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\phi}{ds}\right)^2 - c^2\gamma\left(\frac{dt}{ds}\right)^2 + 1 = 0, \tag{32}$$

with

$$r^2 \frac{d\phi}{ds} = h, \tag{33}$$

and

$$\gamma \frac{dt}{ds} = n - \frac{\epsilon\epsilon'}{c^2 m' r}, \tag{34}$$

where

$$\gamma = 1 - \frac{2\kappa m}{c^2 r} + \frac{\kappa \epsilon^2}{c^4 r^2}. \tag{35}$$

If we eliminate t, s from (32) by means of (33) and (34), and write $u = 1/r$, we obtain as the geometrical equation of the orbits

$$h^2\left(\frac{du}{d\phi}\right)^2 = c^2\left(n - \frac{\epsilon\epsilon'}{c^2 m'} u\right)^2 - (1 + h^2 u^2)\left(1 - \frac{2\kappa m}{c^2} u + \frac{\kappa\epsilon^2}{c^4} u^2\right). \tag{36}$$

We will reserve a detailed investigation of these orbits for a future communication. We may, however, point out that it follows from (36) that u, and therefore r, can be expressed as an elliptic function of ϕ. Since the coefficients of (36) are real and the right-hand side is necessarily positive, this elliptic function has a real period. Hence the orbit is periodic in the sense that r is a periodic function of ϕ, although the period will not in general be 2π. The orbit will, therefore, not be closed, and there will be a rotation of the apse lines. This periodicity of the orbit is due to the fact that we have taken no account of the modification in the field due to the moving particle. A more complete solution, which included the effects of both particles, would probably lead to orbits which would not generally be periodic. This would correspond to the fact pointed out by Einstein, that moving masses imply radiation of gravitational energy.

Einstein seems to infer from this that his law of gravitation must be subject to the restrictions of the Quantum theory. He says* :—

"Gleichwohl müssten die Atome zufolge der inneratomischen Elektronenbewegung nicht nur electromagnetische, sondern auch Gravitationsenergie ausstrahlen, wenn auch in winzigem Betrage. Da dies in Wahrheit in der Natur nicht zutreffen dürfte, so scheint es, dass die Quantentheorie nicht nur die Maxwellsche Elektrodynamik, sondern auch die neue Gravitationstheorie wird modifizieren müssen."

Einstein may be correct in his speculation, but is there not another possibility? An investigation of the motion of a particle of infinitesimal mass in the field of a particle of finite mass has led us to equation (36) and orbits which are always strictly periodic. The complete solution of the problem of two bodies might well lead to an equation corresponding to (36) and to orbits which are not in general periodic, but which *may* in certain circumstances be periodic. These would give the "quantised" orbits. If it should then appear that the non-periodic orbits tended to asymptote down to the periodic orbits, and we could trace the radiation emitted in the process, the whole secret of the Quantum hypothesis would be laid bare. This is, of course, mere speculation, but it may serve to show the extreme importance of obtaining the exact solution of Einstein's equations corresponding to *two* point singularities.

§ 3. *The Conception of a Point Electron.*

We observe from (22) that the effect of the singularity upon the gravitational field is given by the deviation of γ from its value at infinity, namely unity. As we approach the singularity from infinity γ steadily decreases to a minimum at $r = \epsilon^2/mc^2$ and then increases and finally reaches a positive infinity at the singularity. It is equal to its value at infinity when

$$r = \epsilon^2/2mc^2. \tag{37}$$

For an electron we may take $\epsilon = 3 \times 10^{-10}$ C.G.S. electrostatic units, $m = 10^{-27}$ grm. and $c = 3 \times 10^{10}$ cm./sec. and (37) then gives $r = 0{\cdot}5 \times 10^{-13}$ which is of the same order as the usually accepted value of the radius of an electron. It would be easy to attach too much importance to this numerical result. Firstly, it is not clear that the equality of γ, and, therefore, of the $g_{\mu\nu}$ at two different places has any precise physical meaning. It probably depends upon the mode of measuring r, which in this theory is largely at our disposal. Secondly, the third term in (35) represents the effect of electromagnetic energy upon the gravitational field and by equating γ to unity we have obtained a relation between mass and electromagnetic energy. Such a relation is given by the ordinary electromagnetic theory of mass, and in fact as far as we are aware the only determination of the radius of an electron is by means of a formula for the radius in terms of the mass and charge, which differs from (37) only by a numerical factor which depends upon the distribution of charge in the electron. It is, therefore, not surprising that we should obtain a radius of the same order of magnitude. At the same time the present point of view suggests a new meaning for the "radius" of an electron. On the older theory of electromagnetic mass we were forced to regard an electron as a body of small but finite dimensions; it could not be a mere *point* singularity of the field, as the mass of such a point charge would be infinite.

---

* 'Berlin Sitzungsberichte,' 1916, p. 696.

In the same way we cannot have a point singularity on Einstein's theory if we neglect the effects of electric charge, for in this case $\gamma = 1-2\kappa m/c^2 r$ and $\gamma = 0$ for $r = 2\kappa m/c^2$. Hence at $r = 0$, $g_{11} = 0$, $g_{44} = \infty$ and at $r = 2\kappa m/c^2$ $g_{11} = \infty$, $g_{44} = 0$. We have, therefore, not a solution with a single point singularity, but a solution with a point singularity surrounded by a spherical surface of singularity.

On the other hand if we include the effects of charge, we see from (35) that γ has no zeros or infinities other than $r = 0$, if $\epsilon^2/m^2 > \kappa$, *i.e*, $> 0{\cdot}67 \times 10^{-7}$ C.G.S. units. This condition is amply satisfied, both in the case of the electron and the hydrogen nucleus. We may therefore, if we wish, regard these as true point singularities of the field. As we approach such a point very closely γ becomes very large and the field is profoundly modified. As a measure of the dimensions of this region we may take the radius of a sphere on the surface of which the g's have the same value as at infinity. It is this radius which we have calculated above as the radius of the electron.

§ 4. *The Effect of an Electron upon Radiation in its Field.*

According to Einstein's theory, a ray of light passing near to an attracting mass is deflected, and it appears that the deflection would be very great for a ray passing very close to the attracting mass. Indeed, it would seem to be not impossible that a ray which passed sufficiently close to an attracting particle might be so strongly deflected that it would be permanently entrapped by the particle. If this should prove to be a legitimate deduction from Einstein's theory, it might have far-reaching consequences on our views as to the nature of the electron. It will be shown, however, that if we allow for both the charge and the mass of the electron, even though we regard it as a point singularity, no such result is possible. The deflection is always finite. As the distance of nearest approach diminishes, the deflection increases up to a maximum and then diminishes, becomes negative, and finally approaches $-\pi$, so that the ray is returned very nearly along its original direction.

For the path of a ray of light, we have $ds = 0$, which, for a ray in the equatorial plane $\theta = \frac{1}{2}\pi$, gives

$$\gamma c^2 dt^2 = \gamma^{-1} dr^2 + r^2 d\phi^2. \tag{38}$$

We have also to make the integral $\int dt$ stationary for given initial and final points. From (38), we have

$$\int c\,dt = \int \frac{1}{\gamma}\left\{1 + \gamma r^2 \left(\frac{d\phi}{dr}\right)^2\right\}^{\frac{1}{2}} dr.$$

Hence

$$\delta \int c\,dt = \int r^2 \frac{d\phi}{dr}\left\{1 + \gamma r^2 \left(\frac{d\phi}{dr}\right)^2\right\}^{-\frac{1}{2}} \frac{d}{dr}(\delta\phi)\,dr$$

$$= \left| r^2 \frac{d\phi}{dr}\left\{1 + \gamma r^2 \left(\frac{d\phi}{dr}\right)^2\right\}^{-\frac{1}{2}} \delta\phi \right|$$

$$- \int \frac{d}{dr}\left[r^2 \frac{d\phi}{dr}\left\{1 + \gamma r^2 \left(\frac{d\phi}{dr}\right)^2\right\}^{-\frac{1}{2}}\right] \delta\phi\,dr.$$

The expression to be taken between limits vanishes, since $\delta\phi$ vanishes at the extreme points of the path, and the integral vanishes if

$$r^4 \left(\frac{d\phi}{dr}\right)^2 = p^2 \left\{1 + \gamma r^2 \left(\frac{d\phi}{dr}\right)^2\right\}$$

where p^2 is a constant.

Writing $u = 1/r$, and substituting the value of γ from (35), this may be written

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{p^2} - u^2 + \frac{2\kappa m}{c^2} u^3 - \frac{\kappa \epsilon^2}{c^4} u^4, \tag{39}$$

from which it may be shown that p is the length of the perpendicular drawn from the origin on to the asymptotes of the path of the ray.

The right-hand side of (39) is a quartic in u, and hence u may be expressed as an elliptic function of ϕ. This elliptic function could be worked out explicitly if for any reason it should become desirable to investigate the exact form of the rays. The analysis is however heavy, and we can obtain all the information we require by other methods.

Denoting the right-hand side of (39) by $f(u)$, we have

$$f'(u) = -2u\left(1 - \frac{3\kappa m}{c^2} u + \frac{2\kappa\epsilon^2}{c^4} u^2\right).$$

Inserting the values of m and ϵ given in the last section, it will be seen that the quadratic on the right-hand side has no real roots in the case of the electron or the hydrogen nucleus, and, therefore, the only real root of

$f'(u) = 0$ is $u = 0$. Hence $f(u) = 0$ has at most two real roots. It then appears from (39) that, for real rays $p^2 > 0$, so that all real rays have real asymptotes, and it is not possible for light to circle in a closed "light orbit" round the electron. If $p^2 > 0$, it follows that $f(u) = 0$ has one real positive root and one real negative root. Let $u = \lambda$ be the positive root then, if the angle between the asymptotes is 2α, we have

$$\alpha = \int_0^\lambda \left\{\frac{1}{p^2} - u^2 + \frac{2\kappa m}{c^2} u^3 - \frac{\kappa \epsilon^2}{c^4} u^4\right\}^{-\frac{1}{2}} du.$$

Writing $u = \lambda v$, and remembering that $f(\lambda) = 0$, this gives

$$\alpha = \int_0^1 (1-v)^{-\frac{1}{2}} \left\{1 + v - \frac{2\kappa m}{c^2}\lambda(1+v+v^2) + \frac{\kappa\epsilon^2}{c^4}\lambda^2(1+v)(1+v^2)\right\}^{-\frac{1}{2}} dv. \quad (40)$$

For rays which approach the electron very closely, p is very small and λ is very large, and, for sufficiently large values of λ (40), approximates to

$$\alpha \bumpeq \frac{1}{\lambda}\frac{c^2}{\kappa^{\frac{1}{2}}\epsilon}\int_0^1 \frac{dv}{\sqrt{(1-v^4)}},$$

which tends to zero as λ increases. Hence rays which, if undisturbed, would pass very close indeed to the electron, are reflected back practically along their original direction.

For larger values of p, we may conveniently refer back to equation (39). Differentiating with respect to ϕ, we obtain

$$\frac{d^2u}{d\phi^2} + u = \frac{3\kappa m}{c^2}u^2 - \frac{2\kappa\epsilon^2}{c^4}u^3. \quad (41)$$

Allowing for the difference of the units employed, this agrees with the usual equation for the deflection of a ray of light, except for the presence of the last term, which gives the influence of the electric field.

The equation (41) may be integrated by successive approximation in the usual way if u is so small that the terms on the right-hand side are small compared with u. In this way we obtain as the total deflection of the ray

$$\frac{4\kappa m}{c^2 p} - \frac{3\pi\kappa\epsilon^2}{4c^4p^2}. \quad (42)$$

The first term is the well-known Einstein deflection, and the second term shows that the deflection is reduced by the electric field. This approximation will hold for sufficiently large values of p and we have already examined the case of very small values of p. The intermediate cases may be studied by means of (39) if need should arise.

§ 5. *The Influence of the Sun's Electric Field on the Crucial Phenomena of Einstein's Theory.*

In accordance with these results the deflection of a ray of light by the Sun and the predicted displacement of the lines of the Solar spectrum will be influenced by the Sun's electric field.

The ratio of the wave-length of the radiation from a terrestrial atom to that of the radiation from a similar atom in the Sun is, in accordance with the usually accepted theory, given by $\sqrt{\gamma}$. Now it appears from (35) that the effect of any electric field which the Sun may have, whatever its sign, will tend to counteract the effect of the Sun's mass, and we may inquire what strength of field would be necessary to produce exact compensation. For this purpose we must have $\gamma = 1$ at the Sun's surface, or from (35),

$$\epsilon^2 = 2mc^2r.$$

Such a charge on the Sun, if distributed symmetrically, would give at the Sun's surface a field of intensity

$$c\sqrt{\frac{2m}{r^3}} \text{ c.g.s. units} \quad \text{or} \quad c^2\sqrt{\frac{2m}{r^3}}\, 10^{-8} \text{ volts/cm.}$$

The quantity under the square root being of the same order as the Sun's mean density, this requires a potential gradient of the order of 10^{13} volts. per cm., which is far too high to be admissible. The solution to the vexed question of the displacement of the spectrum lines is not to be found in this direction.

In the same way it may be shown that, for any possible value of the Sun's electric field, the second term in (42) is very small compared with the first, and the electric field will not exert any measurable effect on the bending of the rays of light in the gravitational field of the Sun.

On a Physical Theory of Stellar Spectra.

By M. N. SAHA, D.Sc., Lecturer in Physics and Applied Mathematics,
Calcutta University.

(Communicated by Prof. A. Fowler, F.R.S. Received January 18, 1921.)

1. *Introduction.*

The present paper embodies an attempt towards a physical explanation of the ordered gradation in the spectra of stars—a subject in which pioneering work was done by the late Sir Norman Lockyer, but which was worked up with systematic thoroughness at the Harvard College Observatory, under the lead of the late Prof. E. C. Pickering and Miss A. J. Cannon.* During this interval the spectra of more than 100,000 stars have been photographed, classified, and published with full details in the Henry Draper Memorial Catalogue. The most noteworthy facts which have been brought to light from these monumental studies have thus been summarised by H. N. Russell.†

"The spectra of the stars show remarkably few radical differences in type. More than 99 per cent. of them fall into one or the other of the six great groups which during the classic work of the Harvard College Observatory were recognised as of fundamental importance, and received as designations, by the process of the survival of the fittest, the rather arbitrary letters B, A, F, G, K, M. That there should be so few types is noteworthy, but much more remarkable is the fact that they form a continuous series. Every degree of gradation between the typical spectra denoted by B and A may be found in different stars, and the same is true to the end of the series, a fact recognised in the familiar decimal classification, in which B5A, for example, denotes a spectrum half-way between the typical examples B and A. The series is not merely continuous, it is linear. There exists slight difference between the spectra of different stars of the same spectral class, such as Ao, but these relate to minor details. Almost all the stars of the small outstanding minority fall into three other classes (or rather four), denoted by the letters P, O, N, R. Of these, O undoubtedly precedes B at the head of the series, while R and N, which grade one into the other, come probably at its other end, though in this case the transition stages are not clearly worked out."

Russell is of opinion that the principal differences in the stellar spectra arise in the main from variations in a single physical variable in the stellar atmosphere, and many converging lines of argument show that this variable is the temperature of the stellar atmosphere. Table I shows these facts diagrammatically with provisional values of the temperature.

---

* Harvard, 'Annals,' vol. 28, Parts I and II ; vols. 56, 76, and 91.

† 'Nature,' vol. 93, pp. 227, 252, 281.

Table I.

Stellar class.	Typical star.	Secchi's classification.	Temperature.[*]		Remarks.
			Wilsing and Scheiner.	Saha.	
Pb	The Great Orion Nebula	—	15,000 K[†]		
Pc	I.C., 4997	—	30,000 K	—	Gaseous nebulæ with bright lines.
Oa	B.D. + 35°, 4013	Type V, including Wolf-Rayet stars.	23,000	23,000–24,000	
Ob	B.D. + 35°, 4001		—	22,000	
Od	ζPuppis				
Oe	29 Canis Majoris				Henceforth all lines are dark.
Oe5	τCanis Majoris				
Bo	εOrionis	—	20,000	18,000	
B5A	ηTauri	Type I, Helium and hydrogen stars.	14,000	14,000	
Ao	αCanis Majoris		11,000	12,000	
A5F	βTrianguli		9,000		
Fo	αCarinæ	Type II, Yellow-red stars.	7,500	9,000	
F5A	αCanis Minor		6,000		
Go	αAurigæ		5,000	7,000	The sun is a dwarf star of this class.
G5K	αReticuli		4,500		
Ko	αBootis		4,200		
K5M	αTauri		3,200		
Ma	αOrionis	Type III, Red stars.	3,100	5,000	
Md	0Ceti		2,950	4,000	
N	—	Type IV	2,800		
R	—				

[*] (Compiled from the Harvard 'Annals,' vol. 91, p. 5, and Russell's paper, *loc. cit.* Temperatures given under the heading Saha are calculated from the method given in the present paper.)

[†] Buisson and Fabry, 'Journal de Physique,' 1912, p. 472 (calculated according to the limit of interference method).

It is necessary to dwell a little on the physical basis of stellar classification. The earliest astrophysicists classified the stars according to colour; thus Secchi's type I denoted white stars, type II stood for yellow stars, type III for yellow-reddish stars, and type IV for deep red stars. But Lockyer and Pickering found that the varying intensity of particular groups of absorption-lines in stellar spectra was a far better criterion of stellar classification. Table II, compiled from the Harvard 'Annals,' illustrates the principle.

[For the methods by means of which the intensity of a particular line in different stars has been estimated, reference should be made to Harvard 'Annals,' vol. 27, p. 234; vol. 56, p. 56; vol. 91, p. 5. The following is added for the sake of general explanation:—

Numbers which are underlined denote that the line is bright; otherwise the lines are dark. Lines which are barely visible have the number 1 assigned to them. Faint lines are labelled 2. ⊕ denotes a line of which the intensity cannot be obtained from the Harvard 'Annals.' (?) denotes that the intensity is not stated in numbers in the Harvard 'Annals,' but is compiled from descriptions scattered here and there. The symbol M^+ denotes that the line is due to the ionised atom of the element M, *i.e.*, the atom which has lost one electron. Lockyer called them "proto" elements, and the lines were known as "enhanced lines." We shall subsequently show that they are due to the atom which has lost one electron as the result of the high temperature prevailing in the stars, and acquired a net positive charge.

Table II.—Intensity of Stellar Lines.

Element ...... Line .... Series description	Helium. 4471 *2p–4d*	 4713 *2p–4s*	Parhe 4388 *2p–5d*	He^+ 4686 *3d–4f*	He^+ 4542 *4f–9k*	He^{+*} 4860 *4f–8k*	H* 4860 *2p–4d*	Ca 4227 (*g*) 1S–2P	Ca^+ 3934 (*k*) *1s–2p*	Mg^+ 4481 *3d–4f*
Stellar class.										
Pe	—	—	—	⊕	⊕					
Pf	—	—	—	⊕	⊕	⊕				
Oa	0	—	—	⊕	⊕	⊕				
Ob	0	—	—	*100*	*12*	*10*	0			
Oc	1	—	—	40	*3*	*3*	0	—	—	faint
Od	1	—	—	*20*	10	20*	10*	—	2	faint
Oe	15	2	3	*8*	5	25	10	—	2	1
Oe5	15	4	5	5	4	25	20	—	5	1
Bo	15	5	6	2	2	25	25	—	3	2
B2	22	6	10	1	0	35	35	—	4	3
B3	22	6	10	0	0	40	40	—	4	4
B5	10	3	7	—	—	—	60	—	8	7
B8	5	1	3	—	—	—	80	faint	⊕	7 (?)
B9	4	0	1	—	—	—	90	faint	⊕	7
Ao	0	—	0	—	—	—	100	2	10	10
A2	—	—	—	—	—	—	100	4	40	15
A3	—	—	—	—	—	—	90	⊕	70	⊕
A5	—	—	—	—	—	—	70	⊕	80	⊕
Fo	—	—	—	—	—	—	50	⊕	120	⊕
F5	—	—	—	—	—	—	40	15	150 (?)	faint
Go	—	—	—	—	—	—	20	20	200	faint
G5	—	—	—	—	—	—	15	⊕	200 (?)	0
Ko	—	—	—	—	—	—	10	60	150 (?)	0
K5	—	—	—	—	—	—	5	⊕	⊕	
Ma	—	—	—	—	—	—	2	⊕	⊕	
Mb	—	—	—	—	—	—	—	100	faint	
Mc	—	—	—	—	—	—	0	strong	faint	
Md	—	—	—	—	—	—	—	strong	0	

* The line 4860, generally known as H_β, has been shown in the Table both as a hydrogen and a He^+ line. *For reasons see footnote to* Table VII, and the concluding paragraphs of §6.

The Tables show that the lines of an element just appear at a certain stage, rise step by step to a maximum, and disappear at the other end. Thus the He lines just appear at the Ao class, rise to a maximum intensity at the

B3 class, and go out at the Ob class. Similarly for lines of Ca, Mg, H, and groups of nitrogen and oxygen lines not treated here, because their series formulæ are not known.

2. *Physical Processes at High Temperatures.*

The explanation of these phenomena follows from the theory of "Thermal Ionisation and Thermal Radiation" of gaseous elements developed by the author in a number of papers published elsewhere. The following is a summary.*

Up to this time thermodynamics has been confined to the treatment of physical processes like liquefaction and vaporisation, or chemical processes like decomposition or dissociation of molecules into atoms. It has thus carried us up to the stage where all substances are broken up into constituent atoms. But what takes place if a gaseous mass consisting of atoms only be continued to be heated? Hitherto, the ideas on this point have not been very clear. In his interesting theory of internal constitution of stars, Eddington† indeed suggested that in the interior of stars atoms may be, for the most part, broken up into positive nuclei and free electrons—a view which was also supported by Jeans.‡ Following a suggestion by Nernst,§ Eggert‖ tried to calculate the temperature required for tearing off the outer ring of eight electrons from the Fe-atom; he was the first to apply Nernst's theorem of the reaction-isochore (which enables us to calculate the chemical equilibrium of a dissociating system from the vapour-pressure data of the products of decomposition), to the above mentioned process of the breaking up of the Fe-atom. But not only is the supposed process a hypothetical one, but also the energy evolved in the process, which is essential for the calculation, was obtained in a rather artificial manner. Lindemann¶ calculated the breaking up of the H-atom into the nucleus and the electron in the solar atmosphere by using a method almost identical with that of the author, and taking the ionisation-potential = 13·6 volts in accordance with Bohr's theory.

The author has shown that the first effect of increasing temperature will be to tear off the outermost electron from the atomic system, as represented in the equation,

---

* M. N. Saha, "Ionisation in the Solar Chromosphere" (A), 'Phil. Mag.,' October, 1920; "Elements in the Sun" (B), December, 1920; "On the Temperature-Radiation of Gases" (C), February, 1921; "On Electron-Chemistry and its Application to Problems of Radiation and Astrophysics" (D), 'Journ. Ind. Ast. Soc.,' July, 1920.

† Eddington, 'M.N.R.A.S.,' vol. 77, pp. 16 and 596; vol. 79, p. 2.

‡ Jeans, 'M.N.R.A.S.,' vol. 79, p. 319.

§ Nernst, 'Die Theoretische und Experimentellen Grundlagen des Neuen Wärmesatzes,' 1918, p. 154, Chap. 13, concluding paragraph.

‖ Eggert, 'Phys. Zeit.,' December, 1919, p. 570.

¶ Lindemann, 'Phil. Mag.,' December, 1919.

$$\mathrm{Ca} = \mathrm{Ca}^{+} + e - \mathrm{U}_1. \tag{1}$$

The energy of ionisation for a gm.-atom can be exactly calculated (*i.e.*, in a way free from any hypothesis about the arrangement of electrons about the nucleus), from the ionisation-potential V_i, with the aid of the formula,

$$\mathrm{U}_1 = \frac{e\mathrm{V}_i\mathrm{N}}{300\,\mathrm{J}} \text{ calories.} \tag{2}$$

[N = Avogadro-number, J = mechanical equivalent of heat.]

In cases where the ionisation-potential is not known from direct experimental work, it can be calculated from the convergence-frequency, $1s$, of the principal series of the element with the aid of the quantum relation*

$$\mathrm{V}_i = \frac{h(1s)}{e} \,.\, 300 \text{ (in. volts).} \tag{3}$$

There is now very little doubt that the relation is exact. In certain cases the $1s$-term is very large, and, as the principal series lie beyond the Schumann region, it has not been obtained at all. In such cases the $2s$-term is often mistaken for the $1s$-term. Helium is a well-known example. What are usually called the principal series of helium and the so-called par-helium are not $1s$–mp series at all, but $2s$–mp series.† That this is so clear from the associated property that ordinary cool vapour of helium does not absorb the so-called principal series at all, but is quite transparent to them. If these were the true principal series, helium would have absorbed these lines ($\lambda = 20587$, 10830, etc.), just as sodium-vapour absorbs the D_1 and D_2 lines.

Another good example is thallium, discussed by Foote and Mohler 'Phil. Mag.,' vol. 37, p. 33.

At a given temperature and concentration, a definite equilibrium will be established between the proportions of Ca, Ca^{+}, and e, as represented by the Van't Hoff formula of reversible chemical reaction,

$$\mathrm{Ca} \rightleftarrows \mathrm{Ca}^{+} + e. \tag{4}$$

---

* This relation, which is fundamental in the theory of the ionisation-potential, has now been definitely established for a number of elements. For a comprehensive account see McLennan, 'Proc. Phys. Soc. Lond.,' December, 1918. Franck and Hertz, 'Phys. Zeit.,' vol. 20, p. 132 (1919). The relation has been definitely established for :—Hg, Franck and Hertz, Davis and Goucher, 'Phys. Rev.,' vol. 10, p. 101 (1917); Na, Cd, Rb, Zn, Tate and Foote, 'Phil. Mag.,' vol. 36, p. 64 (1918); Mg, Tl, Foote and Mohler, 'Phil. Mag.,' vol. 37, p. 33; Rb, As, Cs, Foote, Rognley, Mohler, 'Phys. Rev.,' vol. 10, p. 59; Pb, Ca, Mohler, Foote, Stimson, 'Phil. Mag.,' vol. 40, p. 73; Cs, Foote and Meggers, 'Phil. Mag.,' vol. 40, p. 80.

† *Vide* Franck and Reiche, 'Zeits. f. Physik,' vol. 1, p. 154 (1920).

The equation of the reaction-isochore for calculating the equilibrium is*

$$\log \frac{x^2}{1-x^2} \mathrm{P} = -\frac{\mathrm{U}_1}{2{\cdot}3\,\mathrm{RT}} + \frac{5}{2} \log \mathrm{T} - 6{\cdot}5. \tag{5}$$

Where x = fraction ionised, P = total pressure.

Calculations for the ionisation of Ca, Sr, Ba, Mg, Na, K, Rb, Cs, H, He, will be found in the papers A and B.

The next point in question is the identification of the ionised elements in the physical systems in which they may occur. Spectroscopically,† it is quite an easy matter, for the ionised elements show a system of lines which is quite different from the lines of the neutral atom. They are generally known as "enhanced lines." The line-system of the neutral atoms, when classified into series, have the Rydberg number $N = (2\pi^2 e^4 m/h^3)$, but the enhanced lines can be grouped into series only with the Rydberg number 4N. This fact has been very satisfactorily established for He^+, and the alkaline earths by Fowler‡ and Lorenser.§ It is an easy corollary from Bohr's theory of spectral emission that such lines should be due to the ionised atom.

For Zn, Cd, and Hg, Paschen and Wolff‖ have discovered a number of lines in the extreme ultra-violet, which are analogous to the lines of the alkaline earths.¶ Fues shows that these lines are due to the ionised atoms of the elements. For other elements, some lines are supposed to be due to the ionised atom, but these have not yet been grouped into series or critically investigated.**

At the end of § 5, a Table of the chief lines of the neutral and ionised elements are given, with their series-position.

3. *Physical Processes Leading to Ionisation—Radiation of the Characteristic Lines of the Neutral Atom.*††

The process of ionisation cannot be an abrupt one, for the electron, in the course of passing off to infinity, can take up an infinite number of stable orbits characterised by the rotational quantum-number n, and the radial quantum number n' [$n = 1, 2, 3, \ldots, n' = 0, 1, 2, \ldots$], according to the Bohr-Sommerfeld* theory of spectral emission. The various quasi-stationary orbits are represented in the following diagram, which is a slightly modified form of one given by Bohr.†

---

* For the deduction of this formula, see Paper A, *loc. cit.*

† For a *résumé* of the early works to detect ionisation in heated gases, see McLennan, 'Roy. Soc. Proc.,' A, vol. 92, p. 591. McLennan's own experiments with Hg, Zn, Cd, Tl-vapour seems to be quite inconclusive.

‡ Fowler, "Bakerian Lecture," 'Phil. Trans.,' vol. 214.

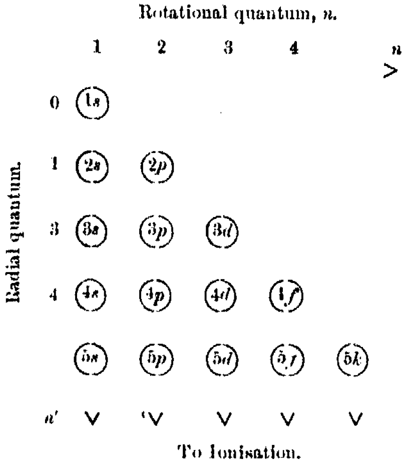
§ Lorenser, 'Inaug. Dissertation,' Tübingen, 1913.

‖ Paschen, 'Ann. der Physik,' vol. 35; Wolff, *loc. cit.*, vol. 42, pp. 825–839.

¶ Fues, 'Ann. der Physik,' vol. 63, No. 17 (1920).

** For example, the Grund-Spektra of Alkalies, discussed by Goldstein, 'Ann. der Physik,' vol. 27, p. 773; Schillinger, 'Wien. Ber.,' p. 608, 1909, and Nelthorpe, 'Astrophysical Journal,' January, 1915; also the enhanced lines discussed by Lockyer and others.

†† See Saha, 'Phil. Mag.,' February, 1921, "On the Temperature Radiation of Gases" (C).



In a mass of atoms, comparatively cool, and not subject to any stimulus, the atoms have their vibrating electron in the 1*s*–orbit. This orbit is circular, corresponding to $n = 1$, $n' = 0$, and has the energy $A-h\,(1s)$. This mass of gas does not emit any light, and when traversed by a continuous beam, can absorb only lines of the principal series, $\nu = 1s-mp$.‡ It cannot absorb lines of the diffuse, sharp or fundamental series. But if we continue to heat up the mass, atoms will progress towards ionisation through the quasi-stationary stages represented in the above diagram, and radiation will follow as a result of the mutual interchange of orbits.§ First of all we shall have emission of 1*s*–2*p* line or lines (fundamental lines), to be followed by the emission of the 1*s*–*mp* series. When the mass of gas emits the line 1*s*–2*p* rather strongly it can absorb lines of the sharp or diffuse series.‖ At a further stage the $2p$-md, or $2p$-ms lines will begin to be emitted. The gas will now be able to absorb the fundamental series. The whole process can thus be symbolically written :—

---

* Sommerfeld, "Atombau und Spektral-Analyse."

† Bohr, 'Zeitschrift für Physik,' vol. 2, p. 434 (1920).

‡ This fact, which comes out as a necessary corollary of the Bohr-Sommerfeld theory, has been experimentally established for Na-vapour (Wood, 'Astr. Jour.,' vol. 29, p. 97; vol. 43, p. 73); for the other alkalies (Bevan, 'Roy. Soc. Proc.,' A, vol. 83, pp. 821–828; vol. 84, pp. 209–225; vol. 85, pp. 54–76), and Hg (Wood, 'Phys. Zeit.,' vol. 14, pp. 191–195 (1913)), for cadmium (Wood, 'Astr. Jour.,' vol. 29, p. 211).

§ For attempts to apply the quantum-theory to these cases, see the following papers:— Einstein, 'Phys. Zeit.,' 1918, p. 124; Planck, 'Ann. der Physik,' vol. 50; 'Berliner Sitz. Ber., 1916; 'Verh. D. Physik. G.,' vol. 17 (1915).

‖ As a matter of fact, the 2*p*–*ms* lines, as well as the 2*p*–*md* lines, have not been reversed at all in the laboratory for any element, when the light is passed through

$$\begin{tikzcd}
1s\text{-orbits} \arrow[r, "\begin{array}{c}\text{Emission}\\ \text{of } 1s\text{-}mp\\ \text{lines.}\end{array}"'] & 2p\text{-orbits} \arrow[r, "\begin{array}{c}\text{Emission}\\ \text{of } 2p\text{-}md\\ \text{lines.}\end{array}"'] & 3d\text{-orbits} \arrow[r, "\begin{array}{c}\text{Emission}\\ \text{of } 3d\text{-}4f\\ \text{lines.}\end{array}"'] & 4f\text{-orbits.} \\
\begin{array}{c}\text{Absorption}\\ \text{of } 1s\text{-}mp\\ \text{lines.}\end{array} & \begin{array}{c}\text{Absorption}\\ \text{of } 2p\text{-}md\\ \text{lines.}\end{array} & \begin{array}{c}\text{Absorption}\\ \text{of } 3d\text{-}4f\\ \text{lines.}\end{array} & \\
{} \arrow[r] & \text{TO IONISATION.} & &
\end{tikzcd}$$

The theory is not yet sufficiently perfect to enable us to calculate the proportion of atoms in the quasi-stationary orbits characterised by the quantum numbers n and n', but some general statements can be made.[*] The "radiation temperature" (*i.e.*, the temperature at which the fundamental line $1s$–$2p$ is emitted, and $2p$–orbits are just commencing to form) will be higher, the higher is the ionisation-potential of the element. Other orbits will be formed at temperatures intermediate between this and the temperature at which ionisation is completed. At any stage, the proportion of atoms distinguished by the quantum numbers (n, n') will be smaller, the larger are the values of n and n'. So that we see that the $1s$–$2p$ line (and $1s$–mp series) will be the first to come out, the last to go, the least amount of element will show them[†] and will be the most intense line all the while. The $2p$–md, $2p$–ms, $2s$–mp lines will come out later, go out earlier, and will be increasingly fainter. The disappearance of $1s$–$2p$ line may be taken as marking the completion of ionisation.

cool vapour (J. J. Thomson, 'Phil. Mag.,' April, 1919). The Balmer series of hydrogen, which may be regarded as the combined sharp and diffuse series of that element, are not reversed when continuous light is passed through ordinary H_2-vapour ; but Ladenburg and Loria ('Verh. der D. Physik. G.,' vol. 10, p. 858 (1908)) showed that they can be obtained as reversed lines when the absorbing column of H_2-vapour is traversed by an electric current. Similarly, Paschen obtained the reversal of the so-called principal series lines of He and Parhe ($\lambda = 10830, 20587$) by passing an electric current through the absorbing column of helium-gas ; ordinary helium-gas is quite transparent to these lines. This shows that the lines are not really principal lines.

[*] Based partly on the works of Planck and Einstein (*loc. cit.*), Planck tries to calculate the distribution of atoms in the different orbits, and Einstein considers the emission and absorption of radiant energy from the standpoint of quantum-statistics. See also Epstein, "Zur Theorie des Starkeffektes," 'Ann. der Physik,' vol. 50, p. 489, who considers the relative probability of passing to the quasi-stationary orbits characterised by the quantum numbers n and n'.

[†] *Cf.* the experiments of Du Gramont on the "Raies Ultimes" of elements, which are either the $1s$–$2p$ line or such lines of $1s$–mp series as lie in the visible region. Du Gramont, 'Bull. Soc. franç. de Phys.,' 1911, No. 14.

4. *Theory of Second-step Ionisation.*

Let us now consider what takes place when the first-step ionisation is completed. Suppose that we have a gaseous mass consisting wholly of Ca^+-atoms, and we continue to heat it up. It is clear that we shall have the emission of the ionised elements, *i.e.*, of the enhanced lines, followed by the commencement of a second-step ionisation, in which Ca^+ loses another electron, and acquires a net double positive charge.

The second-step ionisation is represented by the formula

$$M^+ . \rightleftarrows M^{++} + e - U_2. \qquad (6)$$

The value of U_2 is generally very large, and have not been at all experimentally investigated. But, in the case of helium and the alkaline earths, the enhanced lines have been completely studied and grouped into series. From this we can calculate the energy of second-step ionisation by applying formula (2).

Second-Step Ionisation.*

Element.	1*s*.	Ionisation-potential.	U_2 in calories.
He	4N	53·4 volts.	$12·3 \times 10^5$
Mg	121,270	15·02 „	$3·46 \times 10^5$
Ca	95,740	11·86 „	$2·70 \times 10^5$
Sr	88,840	10·70 „	$2·50 \times 10^5$
Ba	80,575	9·86 „	$2·30 \times 10^5$
Zn	147,544	18·20 „	$4·18 \times 10^6$
Cd	140,226	17·30 „	$3·98 \times 10^5$
Hg	161,800 (?)	19·95 „	$4·60 \times 10^5$

Let us suppose that the fractions x and y of the total Ca-atoms are dissociated to Ca^+ and Ca^{++}-atoms respectively. Then it can be easily shown that the equation of chemical equilibrium is of the form

$$\log P \frac{x^2}{(1-x-y)(1+x+2y)} = -\frac{U_1}{2·3\,RT} + \frac{5}{2} \log T - 6·5, \qquad (7)$$

$$\log P \frac{y^2}{x(1+x+2y)} = -\frac{U_2}{2·3\,RT} + \frac{5}{2} \log T - 6·5. \qquad (8)$$

Since we are confining ourselves to the range where the proportion of neutral Ca-atoms is vanishingly small, we can put $x+y = 1-\epsilon$, where ϵ is a small fraction. The equation (8) then takes the form

---

* Fowler, 'Phil. Trans.,' A, vol. 214; and new calculations not yet published; Fues, 'Ann. d. Physik,' vol. 63, No. 17; Fues's figures for the 1*s*-terms of Zn^+ and Cd^+ are 159,000 and 155,000. The above figures are due to Fowler. The figure for Hg^+ has been calculated by me by using the interesting relation that the 1*s*-term of the enhanced series is roughly 17/9 times the 1*s*-term of the neutral line system in the case of the alkaline earths, and Zn and Cd. *Vide* Fues, *loc. cit.*

$$\log \frac{Py^2}{(1-y)(2+y)} = -\frac{U_2}{2{\cdot}3\,RT}+\frac{5}{2}\log T - 6{\cdot}5. \qquad (8')$$

It is no longer necessary to consider equation (7). The following Table shows the percentage of second-step ionisation of the alkaline earths at pressures of 1 and 10^{-1} atms. Percentage of the first-step ionisation is also shown for the sake of comparison.

Table III.

Element temp.	Mg.	Mg⁺.	Ca.	Ca⁺.	Sr.	Sr⁺.	Ba.	Ba⁺.
10,000	56-90	—	85-99	8-26	90-100	14-40	94-100	23-56
11,000	75-96	—	93	18-44	95	26-64	—	40-78
12,000	86-98	—	97	31-70	98	40-89	—	59-91
13,000	93-99	10-35	99	47-86	—	61-92	—	75-96
14,000	96	21-54	—	64-93	—	76-96	—	86
15,000	98	32-72	—	81-97	—	88	—	92
16,000	99	46-85	—	87	—	92	—	96
17,000	—	61-92	—	92	—	96	—	98
18,000	—	74-96	—	95	—	98		
19,000	—	83-98	—	99				
20,000	—	89-99	—	99				
21,000	—	93						
22,000	—	96						
23,000	—	98						

In each pair of columns, the first gives the percentage of ionisation under a pressure of 1 atm., the second under a pressure of 10^{-1} atms.

The following Table gives the ionisation of helium (both first-step and second-step) at temperatures higher than 16,000° K.

Table IV.—First-Step Ionisation of Helium.

Pressure temperature.	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	Atm.
17,000	9	28	66	94	99	
18,000	14	41	81	97	—	
19,000	21	57	90	99		
20,000	31	71	95	—		
21,000	41	81	98			
22,000	53	89	99			
23,000	64	93	—			Complete.
24,000	73	96				
25,000	81	98				
26,000	87					
27,000	91					
30,000	98					

Table V.—Second-Step Ionisation of Helium.

Pressure temperature.	10^{-3}.	10^{-4}.	10^{-5}.	10^{-6}.
24,000	—	10	24	60
25,000	—	14	39	78
26,000	—	23	57	90
27,000	10	33	72	93
28,000	17	47	86	98
29,000	26	63	95	Complete.
30,000	35	77	97	"

The second-step ionisation of helium is too small even in stars having the highest temperature. Consequently, He^+ lines should be observed even when the temperature reaches 40,000° K., but actually they disappear from the Pc-class of gaseous nebula, the temperature of which is not probably so high. Probably this is an effect of low pressure. According to Emden,* $P = 10^{-6}$ atm. in nebulæ. This brings the temperature of the completion of the ionisation of He^+ to 29,000°–30,000° K., which is not improbable.

Table VI gives the ionisation of hydrogen. It will appear from the Tables that hydrogen should definitely disappear when the temperature reaches the value 22,000° K., but this is not the case, for, according to the 'Harvard Annals,' the Balmer lines occur not only up to the Oa-class, but also in the Pa-class of nebulæ, the temperature of which is probably very large.

Table VI.—Ionisation of Hydrogen.

Pressure temperature.	1-atm.	10^{-1}.	10^{-2}.	10^{-3}.	10^{-4}.	10^{-5}.
7,000	—	—	—	1	4	1·5
7,500	—	—	1	3	8	12
8,000	—	—	2	5	18	26
9,000	—	2	6	20	55	44
10,000	2	5	18	49	86	89
11,000	4	13	39	80	97	
12,000	9	28	68	94		
13,000	16	45	84	98	—	Complete ionisation.
14,000	27	65	93			
15,000	40	81	97			
16,000	55	90				
17,000	69	94				
18,000	80	97				
19,000	87					
20,000	92					
21,000	95					
22,000	97					
23,000	—					

---

* Emden, 'Gaskugeln,' p. 282.

5. *Characteristic Lines of Neutral and Ionised Elements.*

In § 3 we remarked that the ionised elements have a system of characteristic lines which are quite different from the line-system of the neutral element. Table VII shows the chief lines of both systems side by side for a few elements. Here $1s-mp$, $2p-md$, $2p-ms$, $3d-mf$ respectively stand for the principal, the diffuse, the sharp, and the fundamental series. The $1s$ term is generally single, but the other terms are double, treble, or sometimes quadruple. In the principal series m begins with 2, in the diffuse with 3, in the sharp with 2, and in the fundamental with 4. Generally the headline of the series, and in some cases the next one or two, are given.

Table VII.

Element.	Principal $1s-mp$.	Diffuse $2p-md$.	Sharp $2p-ms$.	Fundamental $3d-4f$.	Remarks.
H	1215·3 (etc.).	H_α 6562·8 H_β 4861·3 (etc.).	—	18,751	The sharp series coincides with the diffuse series.
He	570–480 (?)	5875 4471 6678 4922	7065 4713 7281 5047	? ? —	Helium. So-called parhelium.
He^+	304-278 (?)	1640 1215 1087 (etc.).	—	4686	The sharp series coincides with the diffuse series as in H.
Mg	2852·11 4571·38	3838·29 32·31 29·36	5183·62 72·70 67·38		
Mg^+	2795·52 2802·70	2797·99 2790·77	2936·50 2928·68	4481	The 4481 line is also double.
Ca	g 4226·73 6572·78				
Ca^+	K 3933·66 H 3968·47	— —	3736·90 06·02		
Sr	4607·34				
Sr^+	4077·88 4215·66	— —	4305·60 4161·95		
Ba	5535·69				
Ba^+	4554·21 4934·24				
Zn	2139·33 3075·99				
Zn^+	2025·57 2061·98	— —	2558·03 2502·11		
Cd	2288·79 3261·17				
Cd^+	2144·45 2265·13	— —	2748·68 2573·12		
Hg	1849·0 2536·72				
Hg^+	?	—	2847·85 2224·70		

The principal series of hydrogen is given by the formula $\nu = \mathrm{N}(1/1^2-1/m^2)$, the combined sharp and diffuse series by $\nu = \mathrm{N}(1/2^2-1/m^2)$, and the fundamental series by $\nu = \mathrm{N}(1/3^2-1/m^2)$. They are usually named after Lyman, Balmer, and Paschen respectively.

The classification given for helium lines is somewhat different from that usually given. The lines 20857, 5016... and 10830, 3889..., which are usually regarded as the principal series of parhelium and helium, are here regarded as $2s$-mp series; for reasons see § 2. The $1s$-term is, as Franck remarks,* probably single, and is calculated from the ionisation-potential 25·4 volts, as experimentally obtained. The head-lines are calulated by taking $2p$ from the parhelium terms.† Probably 570–480 may be the lines detected photo-electrically by Richardson and Bazonni.‡

The figures for the principal series of He^+ are calculated from the series formula $\nu = 4\,\mathrm{N}(1/1^2-1/m^2)$; the lines, of course, have not been identified. The combined sharp and diffuse series lines $\nu = 4\,\mathrm{N}(1/1^2-1/m^2)$ have been obtained by Lyman. The 4686 is the head-line of the Bergmann or "fundamental" combination $\nu = 4\,\mathrm{N}(1/3^2-1/m^2)$. The next series combination, $4f$-mk-..., is represented by the formula $\nu = 4\,\mathrm{N}(1/4^2-1/m^2)$ and include the Pickering lines $\nu = \mathrm{N}[1/2^2-1/(m+\frac{1}{2})^2]$ as its odd members, while the even members reduce to the Balmer formula for hydrogen, $\nu = \mathrm{N}(1/2^2-1/m^2)$. Their wave-lengths are, however, slightly different, owing to slightly different value of N for He^+. The intensity of these lines is also very small.§ The series $\nu = 4\,\mathrm{N}(1/3^2-1/m^2)$ and $4\,\mathrm{N}(1/4^2-1/m^2)$ were first obtained in the laboratory by Fowler.‖

For the neutral atom of the alkaline earth group and Zn, Cd, Hg the principal lines given are those corresponding to the series combination 1S–2P, $1\mathrm{S}$–$2p_2$.¶ These are the most important lines, for experiments on ionisation-potential show that these are the first to come out when the gas is bombarded by electrons possessing the proper quantum-voltage. The ionised atoms show a system of doublets like the alkalies.

The interest of the above Table lies in the fact that it affords a possibility of spectroscopically testing the ionisation produced by a high temperature. For example, let us take a small quantity of calcium, sealed within a quartz tube which is fitted with plane ends, and then begin to heat the tube, and allow a beam of continuous light to pass through the tube. Then as soon as calcium vapour is formed, the spectrum of the transmitted beam will, on spectroscopic examination, be found to be crossed with the dark absorption lines $\lambda = 4227$, $\lambda = 6573$, and other associated lines of the $1S-mP$, $1S-mp_2$ series. The H, K lines will not be reversed; but if we go on increasing the temperature, then provided that the tube does not collapse, we shall find absorption lines of the diffuse and sharp series lines. Ultimately we shall find the H, K lines as reversal lines. The appearance of these lines may be taken as an indication that a good proportion of Ca atoms has become ionised.

---

* 'Zeits. f. Physik,' vol. 1, p. 154.

† Franck, *loc. cit.*

‡ 'Phil. Mag.,' vol. 34, p. 285 (1917).

§ Evans, 'Phil. Mag.,' vol. 29, p. 284 (1913).

‖ 'Phil. Trans.,' vol. 214.

¶ *Vide* Fowler, *loc. cit.*

On account of the high value of the ionisation-potential of Ca and the elements given above, it is very doubtful, however, whether experiments described above can succeed with quartz tubes or tubes of similar refractory substances known up to the present time. Experiments with the alkali elements Cs, Rb, K are more likely to succeed, for these have very low ionisation-potentials, but the knowledge of their enhanced lines is very meagre.

In view of the total lack of laboratory data to test our hypothesis, we have to turn to the stellar data given in § 1. This is done in the next section.

6. *Physical Processes taking place in Stars; Temperature of Emission of the Stellar Atmospheres.*

The physical meaning of the appearance and disappearance of lines given in § 1, Table II, now becomes apparent. Let us start with the calcium "g" and $Ca^+ - K$ lines.

The "g" line appears in maximum intensity from the very stage when the star begins its effective life. There is no trace of the K line at the lowest stages. This just begins to appear at the Mc-stage, showing that calcium has just begun to be ionised. As we go higher up, the ionisation increases (g is fading off, K becoming more intense). The "g" line completely disappears at the B8-stage, showing that here all calcium has been ionised to Ca^+. The K line reaches a maximum intensity at the G5-stage, then steadily diminishes, showing that a second step ionisation has begun. It disappears completely at the Oc-stage showing that all Ca^+ has been further ionised to Ca^{++}, from above this stage.

Now referring back to Table III, we can assign the following temperatures to the points of completion of ionisation,

B8	Ca completely ionised	13,000° (Ca 1·5 p.c.).
Oc	Ca^+ completely ionised to Ca^{++}	20,000° (Ca^+ 1 p.c.).

and we have also at the stages

Mc	Ca just begins to be ionised	4,000° (Ca^+ just traces).
Go	Maximum proportion of Ca-atoms	70,00°.

Both Ca and Ca^+ can be identified with the aid of their fundamental lines $1s$–$2p$. But such is not the case with any other element excepting Sr, Sr^+, Ba, and Ba^+, for which we have no satisfactory data. It is well known that owing to atmospheric absorption, observations have to be confined to the spectral region $\lambda = 3600$ to about $\lambda = 6000$. It seems to be generally recognised, but nowhere clearly stated, that this fact alone tends to give decided preference to certain elements to the exclusion of others.

For example, let us take hydrogen. In the stars, it can be identified by the $2p$–md lines, the $1s$–mp lines falling within the region of atmospheric absorption. But according to § 2, the $2p$–md lines cannot be absorbed by the ordinary H-atoms, but by such H-atoms of which the electron is in the $2p$ orbit. We can have such orbits only at high temperatures. Hence we find the Balmer lines gradually disappearing in the lower stages. We have at the stage

Mb ...... $2p$-orbits of H just appearing ...... T = 4500°.

On the other side of the scale we find the Balmer lines of hydrogen disappearing at the Ob stage. This is an effect of complete ionisation. We have, therefore, at the stage

Ob ...... Complete ionisation of H...... 22,000° C. (Table VI).

But if some day we can overcome the limitations imposed by the atmospheric absorbing layer, and examine stellar spectra up to 1000 Å.U., it is very probable that we shall get the Lyman lines ($1s$–mp) even at the lowest stage at which a star begins its life,* just like Calcium "g."

The disappearance of Balmer lines from a certain class therefore does not mean that hydrogen is absent from this class, but rather that the stimulus is not sufficiently great to bring out the lines lying within the range of observation.

The position of helium is quite similar. It cannot be identified by its $1s$–mp lines, but by the $2p$–md lines (*vide* Table VII). This can take place only when a sufficient proportion of He-atoms have been converted to the $2p$ state. Owing to the higher value of ionisation-potential of helium, this takes place, as Table II shows, at a much higher stage than that of hydrogen, viz., at the Ao stage. We have, therefore, at the stage

Ao ...... $2p$-orbits of He just appearing ...... T = 12,000° K.

The helium lines are very persistent, as Table II shows. They occur faintly in the Ob and Oa classes, and even in the gaseous nebula† up to the P*d* stage.

---

* Provided, of course, the continuous spectrum from the stellar nucleus extends up to $\lambda = 1000$ Å.U., and is not wholly absorbed by the stellar atmosphere itself.

† Harvard 'Annals,' vol. 91, p. 5.

Here the transition stages have not been satisfactorily worked out. Now a glance at Table IV shows that if we take the pressure 1 atm., helium does not become completely ionised even at a temperature of 30000°. For a pressure of 10^{-1} atm., the temperature of complete ionisation is 25000° K.

The ionisation of helium becomes marked from the B2-stage, as is shown from the appearance of the 4686 line. According to Fowler, this is the $3d$–$4f$ line of He^+, and according to our theory, this requires for its absorption, not only a rather greater stimulus than that which suffices for mere ionisation, but also a greater concentration of the He^+-atoms. We can, therefore, assign to the stage

B2A ... Ionisation of He considerably advanced ... T = 17,000°.

I have sought in vain for satisfactory data for Mg, which can be detected by the $2p$–$3d_j$ lines only (*vide* Table VII). Mg^+ is detected by the line 4481, which, according to Fowler, has the series-combination $3d$–$4f$. A glance at Table II shows that, at the Go-stage, the ionisation of Mg has been considerably advanced, reaching its completion at the Oa-stage. We have, therefore, at the stage

Oa ...... Mg^+ Completely ionised......... T = 23,000° K.

Unfortunately, we have no data of the exact stellar classes from which the Sodium, D_1, and D_2 lines, the Sr line 4607, the Ba line 5543, and the Mg triplet 3838, 3832, 3829, the Sr^+ 4215, 4077, and Ba 4554, 4904, ... just disappear. If these were available, they would have served as good landmarks for determining the temperature of emission of the stellar atmospheres at their various stages of evolution. If these arguments are correct, we are not justified in speaking of a star as a Hydrogen, Helium, or Carbon star, thereby suggesting that these elements form the chief ingredients in the chemical composition of the star. The proper conclusion would be that under the stimulus prevailing in the star, the particular element or elements are excited to radiation of their characteristic lines, while other elements are either ionised, or the stimulus is too weak to excite the lines by which we can detect the elements. This strikes at the root of the prevailing conception of primordial elements.

Table VIII shows the above results at a glance.

A serious discrepancy is shown by hydrogen, which, as shown by the presence of the Balmer lines, is present even beyond the Oa-stage (*vide* Table II, He^+, 4860; we called it an He^+ line there in anticipation of the arguments which follow). But modern spectroscopic work has shown that the lines represented by the Balmer formula $\nu = N(1/2^2 - 1/m^2)$ cannot only be due to hydrogen, but may be due to He^+ and to Li^{++}. If the

Table VIII.

Phenomena.	Stellar class.	Temperature.	Remarks.
Appearance of the K line	Mc	4,000 K	Beginning of the ionisation of Ca.
Disappearance of the "g" line	B8A	13,000	Ca completely ionised.
Appearance of Mg^+ 4481	Go	7,000	Mg considerably ionised.
Disappearance of the K line	Oc	20,000	Ca^+ completely ionised.
Mg^+ 4481 disappears	Oa	23,000	Mg^+ completely ionised.
Appearance of 4686	B2A	17,000	He considerably ionised.
Disappearance of 4471	Oa	24,000 (10^{-1} atm.)	He completely ionised.
Appearance of Balmer lines	Mb	4,500	Appearance of the 2*p*-orbits of H.
Appearance of He lines	Ao	12,000	Appearance of 2*p*-orbits of He.
Maximum absorption of hydrogen lines	Ao	12,000	Maximum concentration of 2*p*-orbits of H.
Maximum absorption of helium lines	B2A	17,000	Maximum concentration of 2*p*-orbits of He.
Disappearance of 4295	B8A	14,000 (10^{-1} atm.)	Sr^+ completely ionised.
Disappearance of Balmer-hydrogen lines	Ob	22,000	H completely ionised.
Disappearance of 4686	Pe	25,000°–30,000°	He^+ completely ionised.

lines be due to He^+, they should have the same intensity as the Pickering lines $\nu = 4N[1/4^2-1/(2m+1^2)]$, for they form the even members of the 4*f*–*mk* series of He^+. Now, if we look back at Table II, we find that this is actually the case from the Oa- to the Oc-class, but in Od and Oe, while the Pickering lines are fading away, the Balmer lines are gaining in intensity.

This fact, taken along with the result that at 20,000° K, 10^{-1} atm. pressure hydrogen is completely ionised, leads us to the conclusion. The Balmer lines due to hydrogen disappear from the Ob-class; those occurring in the Oa-, Ob-class, are due to He^+. In the Oc-, Od-, and Oe-class, there are blends of H and He^+ lines, but below B2A they are entirely due to hydrogen.

The question can be settled by accurately measuring the wave-lengths of the Balmer lines in the different spectral classes, for, owing to slightly different values of N in the case of hydrogen and helium, the wave-lengths are slightly different, as the following Table shows:—*

Hydrogen Lines.	He^+ Lines.
$\nu = N_H[1/2^2-1/m^2]$,	$\nu = 4N_{He}[1/4^2-1/(2m)^2]$.
$M = 1$ $H\alpha$, $\lambda = 6562{\cdot}8$,	$M = 1$, $\lambda = 6560{\cdot}1$.
$M = 2$ $H\beta$, $\lambda = 4861{\cdot}3$,	$M = 2$, $\lambda = 4859{\cdot}3$.
$M = 3$ $H\gamma$, $\lambda = 4340{\cdot}5$,	$M = 3$, $\lambda = 4338{\cdot}7$.
$M = 4$ $H\delta$, $\lambda = 4101{\cdot}7$,	$M = 4$, $\lambda = 4100{\cdot}0$.

---

* Fowler, "Bakerian Lecture," 'Phil. Trans.,' A, vol. 214; Paschen, 'Ann. d. Physik,' vol. 49.

The lines $\nu = \mathrm{N}[1/2^2 - 1/m^2]$ also occur in the highest class of gaseous nebulæ, Pa, which is far higher up than the class from which He$^+$ disappears. There is thus a certain anomaly attached to the Balmer lines.

7. *Concluding Remarks.*

The work thus corroborates Russell's view that the continuous variation of stellar spectral types is mainly due to the varying values of the temperature of the stellar atmosphere, and the classification B, A, F, G, K, M, which has been adopted by the Harvard Astrophysicists, as the result of long years of study and observation, are therefore seen to acquire a new physical significance.

Some minor differences may be noted here. It appears that the temperatures assigned by Wilsing and Scheiner to stellar classes below Go are rather too low. Wilsing and Scheiner assign a temperature of 5000° K. to stars of the Go-class, of which the sun is a typical example. But Biscoe[*] has shown in a comprehensive work that the black-body temperature of the sun is in the neighbourhood of 7500° K. Wilsing and Scheiner's value is therefore too low. The temperatures assigned to the K5, Ma, Md classes (3200°, 3100°, 2950° K.) are too close to each other to explain the large difference in the spectra of these types. Moreover, if these were the true temperatures, the dwarf stars of these classes would possess liquid or solid crusts consisting of carbon, tungsten, tantalum, and probably of some compounds.[†]

From the fact that Ca$^+$ first comes out in considerable quantities in the Mc class, we are probably right in assigning to this class a temperature of 4000° K. But more work is required in the spectral classification of the red stars before we can come to any definite conclusion on this point.

Adams and Kohlschütter[‡] have recently shown that there exists considerable difference in the spectra of the giant and the dwarf stars of the same spectral class, when the intensity of certain groups of lines are compared. It will be seen from Sections 2 and 3 that the nature of the spectra is decided not only by the temperature, but also by the concentration of atoms in the stellar atmosphere. Probably the mass of the star and the average density act differently on elements of different atomic weight, so far as the concentration in the effective layer from which the absorption lines originate is concerned. The problem, however, awaits further investigation.

---

[*] Biscoe, 'Astrophysical Journal,' vol. 46 (1917).

[†] The melting points of tungsten and tantalum are 3800° K. and 3200° K. Carbon is a volatile substance, *i.e.*, has its boiling point lower than the melting point. But Lummer claims to have succeeded in melting carbon by running the arc in an enclosed space under its own vapour pressure. He estimates the melting point to be 4200° K. under a pressure of 1 atm. of its own vapour.

[‡] Adams, 'Communications from Mount Wilson Solar Observatory,' No. 28, 1916.

In spite of the uncertainty in choosing the proper concentration for a given element, it will be admitted from what has gone before that the temperature plays the leading *rôle* in determining the nature of the stellar spectrum. Too much importance must not be attached to the figures given, for the theory is only a first attempt for quantitatively estimating the physical processes taking place at high temperature. We have practically no laboratory data to guide us, but the stellar spectra may be regarded as unfolding to us, in an unbroken sequence, the physical processes succeeding each other as the temperature is continually varied from 3000° K. to 40,000° K.

In conclusion, it is my great pleasure to record my best thanks to Prof. A. Fowler for the interest he has taken in the work, and the many valuable items of information, advice, and criticism with which he has helped me. I also wish to express my thanks to Mr. S. K. Ghosh, M.Sc., of the Calcutta University, for much useful help in the preparation of this paper.

On the Catalytic Dehydrogenation of Alcohols.

By Eric Keightley Rideal.

(Communicated by Sir Wm. Pope, F.R.S. Received February 16, 1921.)

In spite of the numerous industrial processes developed for the hydrogenation of various organic substances with the aid of catalytic materials, but little work has been accomplished on the determination of the mass action equilibrium constants of organic substances which undergo thermal dissociation. Thus in the dissociation of ethyl alcohol into acetaldehyde and hydrogen, or of isopropyl alcohol into acetone and hydrogen, two reactions widely employed, the dependance of the degree of the dissociation on the temperature is unknown. Sabatier ('La Catalyse en Chimie Organique') states (p. 82) that acetaldehyde is easily hydrogenated at 140° C. and acetone at a temperature of 115° to 125° C. utilising nickel as a catalytic agent.

At higher temperatures, however, 200° to 350° C., acetone is not hydrogenated to isopropyl alcohol, but methyl isobutylketone and diisobutylketone are produced. Again (p. 161) ethyl alcohol undergoes rapid dehydrogenation at a copper surface within the temperature range 200° to 350° C.; at 420° C. the acetaldehyde produced undergoes decomposition. Isopropyl alcohol in contact with the same catalytic agent slowly commences to undergo dehydrogenation at 150° C., dehydrogenation being rapid at 250° to 430° C.

It is evident that these reactions are reversible and that the mass action equilibrium constants should be capable of experimental determination if the main reactions proceed smoothly without side reactions occurring to any appreciable extent. According to Sabatier the temperature range suitable for experimental investigation for both these alcohols lies within the range 100° to 350° C.

That nearly complete hydrogenation or dehydrogenation can be effected within such a small temperature range, is confirmed by application of the Nernst heat theorem, utilising the approximation formula. A closer estimation of the dissociation constants of the alcohols by this method is at present impossible, owing to our lack of knowledge of the heats of reaction and the specific heat data.

The approximation formula however indicates clearly that large changes in the values of the dissociation constants are to be expected within a very narrow range of temperature variation and at comparatively low temperatures.

The following values are found in the literature (Landolt Börnstein Tabellen) for the heats of combustion of the various constituents at constant pressure. Owing to lack of information as to the variation of the specific heats with the temperature the values of $\Delta U_{18\,\mathrm{C.}}$ will be taken as equal to $\Delta U_{0\,\mathrm{K.}}$

	Kgm. cal.
H_2	58·1
CH_3CHO	275·5–281·1
CH_3COCH_3	427·2–437·3
C_2H_5OH	326·1–340·5
Isopropyl alcohol	478·7–493·3

The values of the heat of reaction may accordingly vary within wide ranges, thus the value of ΔU_0 for the reaction

$$CH_3CHO + H_2 \rightleftarrows C_2H_5OH + \Delta U_0$$

may vary from +13,900 calories to −6,900 calories according as the upper or lower values for the heat of combustion are taken; the two positive values are 13,900 and 7,500 giving a mean value of 10,700 calories, which will be the one adopted. In the case of the dissociation of isopropyl alcohol there is a similar variation from 16,700 calories to −7,900 calories. The two positive values being 16,700 and 6,700 with a mean of 11,800 calories. On application of the approximation formula of the Nernst heat theorem to the reaction:

$$CH_3CHO + H_2 = C_2H_5OH + 10{,}700,$$

$$K_p = \frac{P_{H_2} \cdot P_{CH_3CHO}}{P_{C_2H_5OH}},$$

$$\log_{10} K_p = -\frac{\Delta U_0}{4{\cdot}571\,T} + 1{\cdot}75\,\Sigma\nu \log_{10} T - \frac{\beta T}{4{\cdot}571} + \Sigma\nu C_0.$$

Taking the chemical constants of the alcohol and aldehyde as equal to one another, the chemical constant of hydrogen as 2·2, and the variation of the specific heats of the reacting organic substances as compensating one another we obtain

$$\log_{10} K_p = -\frac{10{,}700}{4{\cdot}571\,T} + 1{\cdot}75 \log_{10} T - 3{\cdot}75\,.\,10^{-4}\,T + 2{\cdot}2,$$

where β is the variation of the specific heat of hydrogen with the temperature.

Evaluating this expression we obtain

T.° K.	$\log_{10} K_p$.
200	−5·55
300	−1·38
400	+0·63
500	+2·05

Similarly with the reaction

$$CH_3COCH_3 + H_2 \rightleftharpoons CH_3CHOHCH_3 + 11{,}800.$$

Making the same assumptions with respect to the temperature variation of the specific heats and the chemical constants of the ketone and the alcohol we obtain:

T.° K.	$\log_{10} K_p$.
200	−6·75
300	−2·17
400	+0·03
600	+1·78

It is evident, from Sabatier's data, that the values of K, as determined by the Nernst approximation formula, are somewhat too high, although the extremely rapid variation of the constant with the temperature is clearly evident. The variations of the specific heats of the alcohols aldehyde and ketone with the temperature are probably extremely great, since such substances, capable of possessing many degrees of freedom, have, as a rule, a high temperature coefficient for the specific heat, as can be instanced in the case of ether, which has the following values:—

C_v.	Temperature °C.
32	0
32·6	100
41·6	200

Preliminary experiments indicated that the dynamic or flow method of determining the equilibrium was unsuitable, owing to the fact that at low temperatures, even with a very considerable length of catalytic material in a column, unchanged alcohol in large excess was invariably present in the effluent gases. The static method is less objectionable, although side reactions occur, all the more important when the reactants are exposed to the catalytic material for prolonged periods of time. The catalytic activity of various substances for these reactions varies extensively. Thus, with platinum black, and to a less extent with reduced nickel, both the aldehyde and the alcohol undergo partial decomposition relatively quickly:—

$$CH_3CHO \rightleftarrows CH_9 + CO,$$

$$C_2H_5OH \rightleftarrows C_2H_4 + H_2O.$$

The carbon monoxide and ethylene thus produced may undergo still further reactions by hydrogenation, the former to methane and the latter to ethane.

The behaviour of acetone at high temperatures has already been noted; even at low temperatures, acetaldehyde is unstable, undergoing both polymerisation and aldol condensation, with loss of water.

Metallic copper was found to be the most suitable catalytic agent to employ, since its catalytic activity for the reaction of hydrogenation and dehydrogenation is quite marked, whilst disturbing secondary reactions take place relatively slowly on its surface.

It was also found necessary to work with relatively low partial pressures, a closer approximation to the conditions of reversibility being obtained, doubtless associated with the stability of the aldehyde under these conditions.

Although the dissociation of the alcohol to aldehyde and hydrogen takes place relatively rapidly at the commencement, yet the reverse reaction proceeds but slowly, and a long period of contact is necessary even to approximate to equilibrium conditions. According to the general concept of catalytic mechanism, an alcohol molecule striking the surface of the catalyst may either undergo decomposition or it may evaporate from the surface unchanged. For the combination of aldehyde and hydrogen, a molecule of each species has to strike adjacent molecules of the catalytic surface: a much less frequent phenomenon. The following calculation indicates that, unless the number of alcohol molecules striking the surface and evaporating again unchanged is extraordinarily large, the reverse reaction will proceed much more slowly.

If pure alcohol vapour be brought into contact with the catalytic surface, the number of gramme-molecules of alcohol striking each square centimetre of surface per second will be

$$\frac{43{\cdot}75\,.\,10^{-6}}{\sqrt{(\mathrm{MT})}}\,.\,p,$$

where M is the molecular weight of the alcohol, T the temperature in degrees K, and p the pressure in bars.

If, on the other hand, an aldehyde hydrogen mixture be brought into contact with the catalyst, and we assume that an aldehyde molecule occupies a position, X, on the molecular lattice of the catalyst, the chance of combination is dependent on the probability that a molecule of hydrogen will occupy one of the surrounding eight squares. If p' be the partial pressure of the hydrogen of molecular weight M_0, σ the life of a hydrogen molecule on the surface, and d the molecular diameter; then the fraction of the unit area covered with hydrogen will be

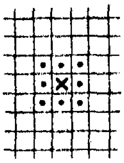


FIG. 1.

$$\frac{43{\cdot}75\,.\,10^{-6}}{\sqrt{(\mathrm{M}_0\mathrm{T})}}\,.\,p'\,.\,\sigma\pi d^2,$$

the probability that a molecule of aldehyde will be in contact with one of the hydrogen will be

$$8\left(\frac{43{\cdot}75\,.\,10^{-6}}{\sqrt{(\mathrm{M}_0\mathrm{T})}}\right)p'\,.\,\sigma\,.\,\pi d^2.$$

This is equal to the total number of contacts divided by the total number of molecules of aldehyde on the surface.

If x be the total number of contacts, p' the partial pressure of the aldehyde, in this case equal to that of the hydrogen, σ' the life of an aldehyde molecule on the surface, M_1 its molecular weight, then

$$\frac{x}{\dfrac{43{\cdot}75\,.\,10^{-6}}{\sqrt{(\mathrm{M}_1\mathrm{T})}}\,.\,p'\sigma'} = 8\left(\frac{43{\cdot}75\,.\,10^{-6}}{\sqrt{(\mathrm{M}_0\mathrm{T})}}\right)p'\pi d^2\sigma,$$

therefore
$$x = 8\left(\frac{43{\cdot}75\,.\,10^{-6}}{\sqrt{(\mathrm{T})}}\right)^2\,.\,\frac{1}{\sqrt{(\mathrm{M}_0\mathrm{M}')}}\,.\,p'^2\,.\,\pi d^2\,.\,\sigma\sigma'.$$

Whilst for the unimolecular alcohol decomposition the rate of reaction was found to be

$$\frac{43{\cdot}75 \,.\, 10^{-6}}{\sqrt{(\mathrm{MT})}} \,.\, p\mathrm{X},$$

where X is the fraction of the molecules striking which undergo decomposition.

The ratio of the two velocities is accordingly given by

$$\frac{\text{v. hydrogenation}}{\text{v. dehydrogenation}} = \frac{8\left(\dfrac{43{\cdot}75 \,.\, 10^{-6}}{\sqrt{(\mathrm{T})}}\right)^2 \dfrac{1}{\sqrt{(\mathrm{M_0M_1})}} \,.\, p'^2\pi d^2\sigma\sigma'}{\dfrac{43{\cdot}75 \,.\, 10^{-6}}{\sqrt{(\mathrm{MT})}} p \,.\, \mathrm{X}}$$

$$= \frac{3{\cdot}5 \,.\, 10^{-4} \,.\, p_1{}^2\pi d^2\sigma'\sigma\sqrt{(\mathrm{M})}}{\sqrt{(\mathrm{M_0M_1T})} \,.\, p \,.\, \mathrm{X}}.$$

According to Langmuir* the average life of a molecule on the surface of a solid is given by the expression

$$\sigma = 22860\, a\sqrt{(\mathrm{MT})},$$

where a is determined approximately from the empiric Freundlich isotherm

$$q = ap^{1/n}.$$

If a value of $5{\cdot}10^{-8}$ cm. is adopted for d, the molecular diameter, T as 500° K., $p' = 0{\cdot}5$ megabar, and $p = 1{\cdot}0$ megabar, the ratio of the two velocities will be

$$\frac{\text{v. hydrogenation}}{\text{v. dehydrogenation}} = \frac{358{\cdot}6 \,.\, 10^{-6} a_1a_2\sqrt{(\mathrm{MT})}}{\mathrm{X}} = \frac{5{\cdot}45 \,.\, 10^{-2} a_1a_2}{\mathrm{X}}.$$

The values of a_1 and a_2 for hydrogen and acetaldehyde respectively at this temperature are probably very small and that for hydrogen considerably less than that for the aldehyde. If arbitrary values of $a_1 = 0{\cdot}01$ and $a_2 = 0{\cdot}1$ are adopted, the velocity of decomposition will be no less than 20,000 times the velocity of hydrogenation if the value of $\mathrm{X} = 1$. Elevation of both temperature and pressure reduces the disparity between the two reaction rates.

The general mode of procedure accordingly was to determine the dissociation of the alcohol in contact with metallic copper under reduced pressures, the constant volume method being adopted. When dissociation was nearly complete, the temperature was allowed to fall and the degree of combination ascertained. Under conditions of true reversibility the two curves plotted from the results should naturally coincide. It was found in general that slight decomposition could not be avoided, and that, although the decomposition curves could be repeated, the values obtained for the process of hydrogenation were by no means so regular. If sufficient time were given for the system to arrive at equilibrium, secondary decomposition occurred relatively rapidly at high temperatures, and after two or three hours' contact at low temperatures.

---

* 'J. Amer. Chem. Soc.,' vol. 40, p. 1390 (1918).

Experimental.

Short pieces of copper wire were repeatedly oxidised and reduced, first in hydrogen and subsequently in ethyl alcohol vapour at continually decreasing temperatures until the surface of the metal underwent spontaneous oxidation on exposure to air at room temperature, the final reduction being performed when the catalyst was in place in the apparatus. The apparatus consisted of a simple constant-volume gas thermometer, the bulb being a piece of glass tube, 15 cm. long and 1·5 cm. diameter, containing 5 grm. of the reduced catalytic copper. The bulb was connected by means of a piece of thick-walled tube to a mercury manometer, provision being made by means of a T-piece and two-way tap, for the evacuation of the bulb and for the admission of the alcohol.

The reaction vessel was heated in an electric furnace, sufficiently long as to ensure the uniformity of temperature over the whole length of the tube.

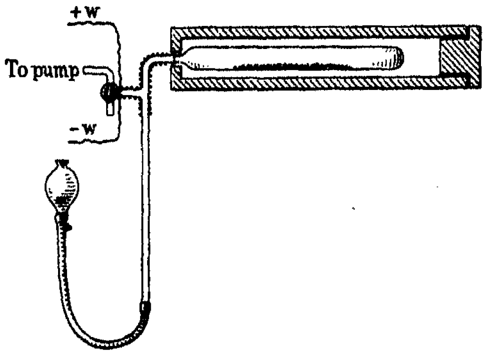


Fig. 2.

To prevent condensation occurring in the system on the top of the mercury in the manometer, the exposed portion of the thin tube was wound with nichrome wire and covered with asbestos, the temperature being maintained, by means of an electric current, at 100° C. during the period of each experiment.

For each experiment the system was evacuated by means of a Sprengel mercury pump, a small quantity of alcohol admitted, and the catalyst freshly reduced at a temperature of 250° C. The tube was then re-evacuated and the temperature maintained at 275° C. to ensure removal of absorbed and dissolved gases.

After cooling a small quantity of the alcohol was admitted and the temperature pressure curve determined at constant volume, time being given for the system to come to equilibrium. Four to five hours were found necessary at the lower temperatures and half an hour at the higher.

The temperature pressure curve was likewise determined during the process of cooling. Long periods of time were generally necessary for the pressure to attain a contant value and it was generally observed that slight secondary decomposition was unavoidable.

Both the ethyl and isopropyl alcohols were pure and free from water.*

As typical of the variations of pressure with the temperature the following table may be cited.

Temperature °C.	Ethyl alcohol.		Isobutyl alcohol.	
	Pressure in mm. Hg.		Pressure in mm. Hg.	
	Heating.	Cooling.	Heating.	Cooling.
20	36	—	18	—
50	64	—	42	—
75	74	75	60	74
100	81·5	83	100	105
125	93	94	119	120
150	102	106·3	145	158
175	114	125	180	204
200	128	148	218	240
225	149	177	250	262
250	185	220	272	277
275	242	261	289	289
300	335	335	—	—

It will be noted that a secondary irreversible decomposition had taken place to a limited extent with each alcohol, since the pressures on the cooling curve are uniformly higher than those on the heating curve. From these data the percentage decomposition at the respective partial pressures was determined, it being assumed that complete vaporisation of the alcohol was effected at 100° C., and at this temperature and pressure the simple gas laws were valid.

If x be the degree of dissociation at any definite temperature and at a pressure of p mm. of mercury the dissociation constant for one atmosphere is obtained from the expression

$$K_p = \frac{x^2}{1-x}\frac{p}{760}.$$

In the following table are given the values of K_p determined from the mean of three experiments on each alcohol, the heating curve alone being utilised for the purpose of calculation. Appended also are given the percentage decompositions of each alcohol under a total pressure of one atmosphere

---

* I am indebted to Mrs. W. G. Palmer for having kindly provided me with carefully purified specimens of these materials.

Temperature °C.	Ethyl alcohol.		Isopropyl alcohol.	
	K_p.	Per cent. decomposition at 1 atmosphere.	K_p.	Per cent. decomposition at 1 atmosphere.
105	$6 \cdot 44 . 10^{-4}$	2·5	$2 \cdot 82 . 10^{-3}$	5·3
150	$1 \cdot 60 . 10^{-3}$	3·9	$2 \cdot 04 . 10^{-2}$	13·6
175	$4 \cdot 77 . 10^{-3}$	6·7	0·177	29·0
200	$1 \cdot 21 . 10^{-2}$	10·6	0·523	51·1
225	$4 \cdot 52 . 10^{-2}$	19·5	1·35	66·9
250	$2 \cdot 32 . 10^{-1}$	38·0	4·82	85·0
275	0·91	60·2	12·0	92·8

If the values of ΔU_0 be calculated for each 25° C. rise by means of the Van t'Hoff equation, the heat of reaction for the dissociation of ethyl alcohol is found to rise from 12,100 calories at 125° C. to 30,500 calories per gramme-molecule at 275° C., whilst for the same temperature range the values for the dissociation of isopropyl alcohol are 25,200 and 20,000 calories per gramme-molecule respectively: this latter substance yielding a mean value of 21,300 calories per gramme-molecule.

In the case of ethyl alcohol the high values obtained are probably due to secondary decomposition of the aldehyde, which is much less stable than acetone.

The value of ΔU_0 obtained from the value of $\Delta U_{373} = 12{,}100$ and from the specific heats of the reactants and products.

$$\frac{d\Delta U}{dT} = C_1 - C_2 = 15 \cdot 80 + 6 \cdot 82 - 20 \cdot 84 = 1 \cdot 78,$$

is 11,500 calories per gramme molecule for ethyl alcohol.

For isopropyl alcohol, taking the specific heat of isopropyl as 27·80 we obtain

$$\frac{d\Delta U}{dT} = 23 \cdot 95 + 6 \cdot 82 - 27 \cdot 80 = 2 \cdot 97,$$

and $\Delta U_{373} = 21{,}300$ or $\Delta U_0 = 20{,}100$.

The mean value of ΔU_0 determined from the heat of combustion of ethyl alcohol was 10,700 calories and from the dissociation curve 11,500 calories. If this figure were corrected for the effect of the alteration of the specific heats with the temperature, on which data are not available, still closer agreement might be obtained.

In the case of isopropyl alcohol there is a very large discrepancy between the mean value of 21,300 calories obtained from the dissociation curve and the mean value of 11,800 calories obtained from data on the combustion of isopropyl alcohol and acetone. The maximum value of 16,700 calories per gramme-molecule is even exceeded.

The heats of combustion of these substances are, however, so large that an error of less than 1 per cent. in the determination would more than compensate for the observed difference. A variation of nearly 3 per cent. in these determinations is actually recorded.

Summary.

In confirmation of Sabatier's experiments the application of the Nernst approximation formula indicates that the variation of the dissociation constants of the equilibria

$$CH_3CH_0 + H_2 \rightleftharpoons C_2H_5OH, \qquad CH_3CO{\cdot}CH_3 + H_2 \rightleftharpoons CH_3CHOHCH_3,$$

with the temperature are extremely large. The velocity of decomposition at the surface of a solid catalyst is shown to be very much higher than the reverse reaction of hydrogenation. The variation of the values of the dissociation constants with the temperature was experimentally determined with a constant volume thermometer containing carefully reduced copper as catalytic material. Concordant values could not be obtained at atmospheric pressures, but a closer approximation to reversibility was obtained when low pressures were employed. The value of ΔU_0 for the decomposition of ethyl alcohol was found to be 11,500 calories, whilst the mean value determined from the heats of combustion is 10,700. In the case of isopropyl alcohol the respective values for ΔU_0 are 21,300 and 11,800. In the latter case an error of less than 1 per cent. in the determination of the heat of combustion of the alcohol would account for the discrepancy.

*On the Design of Diaphragms Capable of Continuous Tuning.**

By Louis Vessot King, M.A. (Cantab.), D.Sc., Macdonald Professor of Physics, McGill University, Montreal.

(Communicated by Prof. A. S. Eve, F.R.S. Received November 11, 1920.)

Section 1.—*Introduction.*

In calculating the acoustic characteristics of an electrically operated submarine sound-generator fitted with a spring-controlled diaphragm, it was found that the apparatus would not radiate efficiently at a given power input and frequency, unless the effective diameter was, within fairly narrow limits, of the correct size, and unless the fundamental pitch was very close to the frequency of the alternating current supplied. The calculations also showed that a non-periodic diaphragm (*i.e.*, one in which there is no elastic restoring force) could not be set into vibration at frequencies in the neighbourhood of 500 cycles without incurring very high copper losses. A similar conclusion was reached with regard to the aperiodic diaphragm. It was concluded that the best efficiency could be secured by operating the sound-generator with alternating current of the same frequency as that of the diaphragm.

Similarly, in most forms of sound-receiving apparatus, the response to submarine sound-waves is most marked when the size of the diaphragm is correctly chosen and when its fundamental pitch is in resonance with the frequency of the incident waves. In the field of submarine acoustic signalling it is evident that a great advantage is to be derived by utilising a means of continuously tuning the diaphragms of the sound-receivers or generators over a certain range of frequencies, thus realising selective transmission and reception. At the same time advantage is taken of the great increase in sensitivity characteristic of highly selective receivers, in which the damping may be reduced to a minimum in view of the possibility of being able to "tune in" to incoming waves as is now done in radiotelegraphy.

Section 2.—*On the Design of Pressure-tuned Diaphragms.*

In order to meet the need for a means of continuously altering the pitch of diaphragms forming part of submarine sound-generators or receivers, the method of construction shown in fig. 1 was adopted. The diaphragm is an accurately turned solid of revolution, preferably fashioned from a single piece

---

* These researches were carried out in June and July, 1917, in connection with researches on submarine acoustics carried out for the Board of Invention and Research, with whose permission the following account is now published. Further details of applications to submarine sound-generating and receiving apparatus will be found in the British Patent Specification 131,041 (1918), "Improvements in and Relating to the Tuning of Diaphragms for Generating or Receiving Sound Waves."

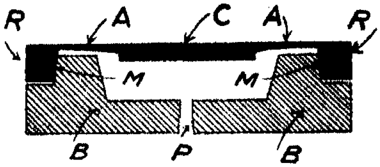


FIG. 1.—Tunable Diaphragm.

M, bronze body ; C, rigid centre ; A, annulus ; R, rigid rim of diaphragm ; M, accurately ground mounting.

of metal or other elastic substance having a low elastic hysteresis and as high an elastic limit as possible. It consists of a solid central portion, C, and of a heavy rim, R, connected by a relatively thin annulus, A. The relative dimensions of C and A may be calculated to give any frequency required in practice, as well as to provide for the maximum range of tuning possible for a material of given elastic properties. The rim of the diaphragm fits very accurately on a slightly tapered mounting of the body, B. The form of the latter may be varied to suit the conditions under which the diaphragm is to be employed. By the application of gaseous pressure (or suction) over the interior of the diaphragm (supposed to be in gas-tight connection with the body), it is caused to bulge slightly (outwards or inwards). As the solid centre is practically undeformed by its displacement relative to the rim, the distance between the outer edge of the former and the inner edge of the latter is increased (or diminished) by an amount depending on the magnitude of the pressure or suction applied. The tension of the thin annulus is thus increased or diminished and as a result the frequency of the vibrations of the central part of the diaphragm is raised or lowered. By this means the fundamental pitch of the diaphragm may be raised or lowered continuously between certain limits by an alteration of gaseous pressure.

The range of tuning which can be effected in this way will depend not only on the design of the diaphragm, but on the material employed in its construction and on the method of mounting. It is evident that a limit is set to the tension allowable in the annulus as soon as the stress-intensity reaches the elastic limit for the material. It is shown by an approximate mathematical investigation that this occurs initially over the inner boundary of the annulus. By suitably varying the thickness of the annulus so as to give a "diaphragm of uniform strength" the radial stress-intensity may be equalised throughout the material, and for a given pitch and elastic limit, the range of tuning may be considerably increased. It was also found that the range of tuning possible within the elastic limits of the diaphragm material varied very slowly with the ratio of radii of the annulus. This conclusion was also verified by experiment and a ratio 2 : 1 finally adopted as most convenient in practice.

Below are given predetermined characteristics of a 3-inch microphone diaphragm:

CHARACTERISTICS OF 3-INCH MICROPHONE DIAPHRAGM.

Inner radius of annulus 1·75 cm., outer radius 3·50 cm.; diaphragm shaped to uniform strength and fundamental pitch of 450; thickness of annulus at inner radius 0·038 cm., at outer radius 0·024 cm.

Mild Steel.—Young's modulus 30×10^6 lbs./sq. in.; elastic limit 50,000 lbs./sq. in.; theoretical frequency-range 450–1210.

Nickel-Chrome Steel.—Young's modulus 30×10^6 lbs./sq. in.; elastic limit 150,000 lbs./sq. in.; theoretical frequency-range 450–1980.

The above values for frequency-range are based on the supposition that the rim and solid centre are absolutely rigid and unyielding. In practice, it is to be expected that the actual limits attained will fall somewhat short of those calculated theoretically, owing to a sensible yielding at the circumference of the diaphragm The importance of securing as closely fitting and as rigid mountings as possible is thus evident: otherwise a large bulge will be required to increase the tension in the annulus to the required degree, resulting in overstress through the effect of large bending stresses not contemplated by theory.

Section 3.—*Experimental Tests of Method of Tuning Diaphragms.*

In order to test the method of tuning diaphragms by pressure or suction, it was decided to embody the invention as part of a simple microphone receiver adapted to the needs of ordinary navigation for receiving signals from submarine sound-generators. Many of these are now established in American and Canadian waters as submarine fog-alarms, and it is probable that their use for this purpose will largely increase in the future. The pitch of these sound-generators is about 500. As it is hardly to be expected that the pitch of such generators can be maintained uniformly constant, it is evident that the receiving microphones, to be most efficient (as well as the telephone receivers used with them) should be capable of continuous tuning over a limited range in the neighbourhood of this frequency.

During a series of tests carried out in July and August, 1917, several types of diaphragms and methods of mounting were tried, with more or less successful result, in the matter of tuning. To obtain the best results it is found essential to avoid the use of rubber or cement packing, and even of soldered contacts. Finally, the type of diaphragm shown in fig. 1, embodied as part of the microphone receiver as shown in fig. 2, was found to give satisfactory results. The heavy rim of the diaphragm was carefully ground to fit the bronze mounting. It is easily seen that pressure or suction tends to bulge the central solid portion of the diaphragm inwards or outwards, thus increasing the tension in the thin annulus. As a result, the stresses produced in the rim of the diaphragm cause the former to grip the mounting more tightly as the pressure or suction is increased. If the taper of the mounting is kept sufficiently small (about 10°), the friction brought into play is sufficient to prevent the diaphragm being forced off by pressure.

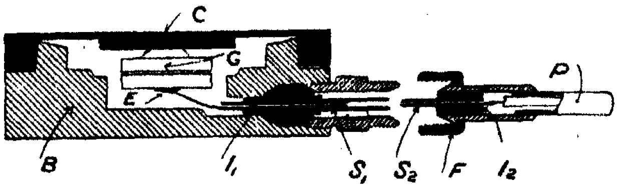


FIG. 2.—Submarine Microphone Receiver fitted with Tunable Diaphragm.

G, carbon granule microphone button ; E, platinized contact to bayonet socket S_1, S_2, mounted in insulating bushings I_1, I_2 ; F, watertight coupling connecting copper tube, P, to microphone body, B. An insulated conductor is passed through the tube and connected through S_2, S_1, E to the microphone button, G. The copper tube itself is the return circuit.

The frequency tests were carried out in a cylindrical tank (57 cm. in diameter) containing water to the depth of 1 metre. To generate sound waves in the water, an ordinary bi-pole receiver was made tight by cementing the diaphragm in position. This was connected to a source of alternating current known as the " Vreeland Oscillator." The alternating current supplied by this apparatus to a circuit free from iron-cored inductances is known to be free from harmonics, and was found especially suited to the work in hand. The frequency was varied over the range 350–1600 by inserting suitable inductances and capacities in the main circuit of the oscillator, and was measured by a specially accurate sonometer method.* By means of a telephone magnet of high resistance connected across the terminals of the apparatus, a steel wire under known tension could be adjusted in length to vibrate in resonance with the frequency of the periodic variations of magnetic flux. The wire was specially mounted to avoid frictional losses at its extremities when vibrating, so that the resonance was extremely sharply defined and determinate. By means of this precision " frequency meter," no difficulty was met with in determining frequencies over the range 350–1600 to an accuracy of one part in a thousand.

---

* This precision frequency meter is described in greater detail in a paper by the writer "On the Determination of the Electrical and Acoustic Characteristics of Telephone Receivers," 'Journal of the Franklin Institute,' pp. 616–617 (May, 1919).

The microphone was immersed to a depth of 38 cm., and was connected to a set of dry batteries and to a high-resistance telephone head-piece in the usual way. By means of an air-pump connected to a steel reservoir and mercury manometer, known pressures could be applied to the microphone diaphragms through the copper tube forming one branch of the electrical circuit. The return circuit, consisting of a rubber insulated wire threaded loosely through the copper tube, was completely protected in this way from inductive effects and "strays," which are always troublesome when listening for very faint sounds.

To obtain the variation of resonance frequency with pressure, the inductances connected to the Vreeland Oscillator were adjusted to produce a note of the desired frequency in the tank. The air-pressure or suction on the diaphragm was then slowly altered until the note heard in the telephone head-piece was a maximum. The pressure giving resonance could be adjusted so that a variation of 2 mm. of mercury gave rise to a noticeable change of audibility. The curve connecting the pressure (corrected for

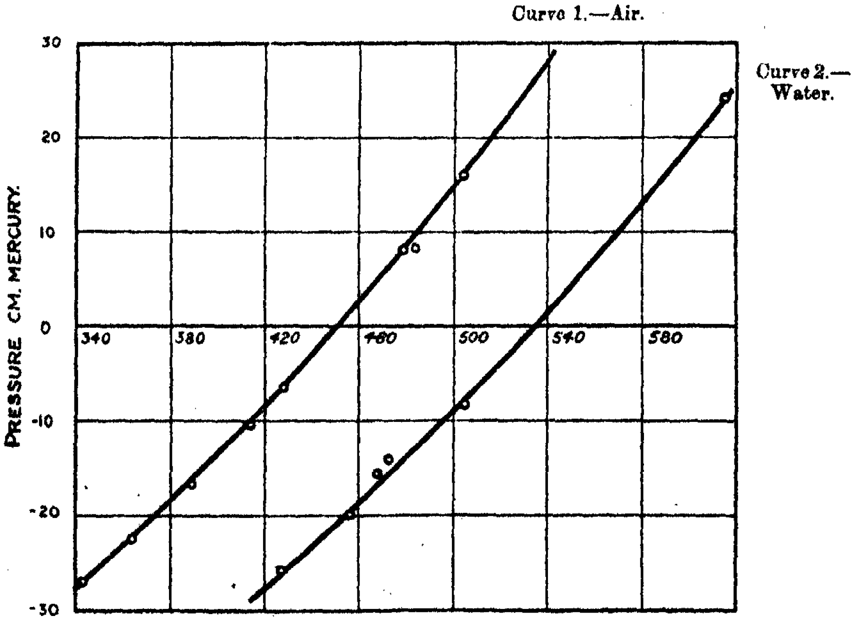


Fig. 3.—Tuning Curve.

Three-inch Tunable Diaphragm. Variation of resonance frequency with applied pressure. Curve 1, diaphragm vibrating in air; Curve 2, diaphragm vibrating in water.

depth of immersion) with resonance-frequency will be referred to as the "tuning-curve" of the diaphragm employed.

A similar curve was obtained for the diaphragm vibrating in air. In this instance the vibrations were excited by a telephone-magnet placed at a few millimetres distance from the steel diaphragm. The two sets of curves thus obtained are shown in fig. 3. It will be noticed from Curve 1 that it is possible to tune from 320 to 480 by variations of pressure within the range +10 cm. to −30 cm., which is easily furnished by the lungs alone.

It was found that the diaphragm showed signs of over-stress for pressures in the neighbourhood of ±35 cm. mercury. As the pressures approach these limits the tuning became less sharp, so that working pressures were subsequently kept well within this range. From previous tests it was ascertained that overstressing a diaphragm did not in any way impair its practical usefulness. The tuning curve was simply shifted parallel to itself, indicating that the effective elasticity of the diaphragm (at zero pressure) had been changed. The sensitiveness of the diaphragm did not appear to have suffered; in fact, diaphragms are most easily adjusted to a required tuning curve by overstressing.

Section 4.—*Note on the Experimental Determinations of the Effective Constants of Diaphragms.*

The shape of the diaphragm is designed to reduce to a minimum the part played by the excitation of overtones. Consider the diaphragm as an elastic system having a single degree of freedom. Denote by x the displacement of the solid central portion of the diaphragm from the zero position resulting from the application of pressure. In a position of equilibrium (under pressure) the elastic restoring force required to balance the applied pressure may be represented by an expression of the form

$$F = \mu_0 x + \lambda x^3. \tag{1}$$

Here μ_0 is a constant depending on the elastic moduli and on the dimensions of the annulus, and is independent of the applied pressure. The term λx^3 is a third order term, in which λ depends on the applied pressure. For a small displacement, ξ, about the new position of equilibrium, the effective restoring force is

$$\delta F = (\mu_0 + 3\lambda x^2)\,\delta x, \quad \text{or} \quad \delta F = \mu\xi,$$

where $\mu = \mu_0 + 3\lambda x^2$ and δx is identified with ξ. We may thus write for the equation of motion representing the free vibrations

$$m\ddot{\xi} + \kappa\dot{\xi} + \mu\xi = 0. \tag{2}$$

The elastic restoring force, μ, and its variation with pressure may be approximately determined by observing a number of stress-strain curves relative to various equilibrium positions. It may also be predetermined roughly from the elastic theory of such diaphragms. The constant, κ, depends (i) on the energy absorbed at the microphone contact; (ii) on that dissipated at the diaphragm mounting, and by elastic hysteresis losses in the material of the diaphragm; and (iii) on the rate of acoustic energy radiation. The magnitude of κ cannot be predicted with certainty, but to obtain high sensitivity and sharp tuning it is evident that all the above factors, except that depending on acoustic radiation, should be reduced to a minimum. To obtain its value experimentally it would be necessary to examine the rate of damping of the free vibrations under various conditions.

The effective mass, m, includes not only the inertia of the vibrating diaphragm itself, but also the "acoustic mass" arising from the inertia of the water set into oscillatory motion by the vibrations of the diaphragm.[*] Considerable information relating to these constants may be obtained from a comparison of the tuning curves obtained in water and air, the acoustic mass in the latter medium being negligible. As long as the damping in water is not too great, the frequency, f_w, is given very approximately by the formula

$$2\pi f_w = \sqrt{(\mu/m_w)}, \tag{3}$$

If we compare f_w with the frequency, f_a, in air when the microphone diaphragm is subjected to the same pressure, so that μ is unaltered, we have

$$2\pi f_a = \sqrt{(\mu/m_a)}. \tag{4}$$

From (3) and (4) it follows that

$$m_w/m_a = (f_a/f_w)^2. \tag{5}$$

If m_w' denote the *acoustic mass* in water, and m_0 the effective diaphragm mass, we have (since the acoustic mass of air is negligible)

$$m_a = m_0, \qquad m_w = m_0 + m_w'. \tag{6}$$

From (5) and (6), it follows that the ratio of the acoustic to the diaphragm mass may be obtained in terms of the resonance frequency ratios corresponding to equal pressures on the diaphragm. If, in addition, μ is known experimentally, and also the resonance frequency in air, (4) and (6) show that m_0, and therefore m_w', may be calculated. For a piston diaphragm of radius a, vibrating in an infinitely extended medium, the acoustic mass is calculated to be

---

[*] *Added Feb.* 28, 1921.—On the theoretical calculation of the acoustic mass of a circular diaphragm of *uniform* thickness filling an aperture in a plane wall which is in contact on one side with an otherwise unlimited mass of fluid, see a recent paper by Lamb, H., "On the Vibrations of an Elastic Plate in Contact with Water," 'Roy. Soc. Proc.,' A, vol. 98, pp. 205–216 (1920). Calculation shows that the theory is unfortunately not applicable to the tunable diaphragms with rigid centres discussed in the present paper.

$$m_w' = \tfrac{8}{3}a^3\rho_w, \tag{7}$$

where ρ_w is the density of the medium. From this equation it is possible to determine the "effective radius" of the diaphragm. As the contribution of the acoustic radiation to the damping constant of a piston diaphragm is known to be

$$\kappa_w' = \tfrac{1}{4}\rho_w\pi a^4 n^2/c, \tag{8}$$

where $n = 2\pi f_w$, and c is the velocity of sound in water, we are enabled in this way to separate out the damping due to acoustic radiation from that due to internal causes.

That the acoustic mass of a small 3-inch diaphragm is comparable to its actual mass, even when vibrating in a confined mass of water, is seen from the following Table, derived from the tuning curves of fig. 3.

Table I.

Pressure in cm. mercury …………	−20	−10	0	10	20
Resonance frequency in air, f_a ……	456	498	536	570	602
Resonance frequency in water, f_w…	375	415	450	487	515
Ratio, $m_w/m_a = (f_a/f_w)^2$ …………	1·48	1·44	1·40	1·37	1·37

The numbers in the last row are seen to be in tolerable agreement, and show that in this instance the acoustic mass is as much as 40 per cent. of the diaphragm mass. The progressive variation of these numbers is attributed by the writer to the influence of the microphone contact employed. The brass mounting of one of the carbon buttons was rigidly screwed into the solid centre, and was insulated from the brass chamber containing the carbon granules and the remaining carbon button by a mica diaphragm. As a result, the steel diaphragm and brass chamber formed an elastically coupled system, with the result that the apparent mass of the former might be expected to vary with the frequency.

Section 5.—*Note on Conditions to be Realised to obtain Maximum Sensitiveness and Sharp Tuning.*

Through being able to continuously adjust the frequency of a diaphragm so as to be able to "tune in" to a note of given pitch, high sensitivity and sharp tuning are important objects to be aimed at in the design of submarine receiving apparatus. As is well known from the theory of the vibrations of damped systems, high sensitivity and sharp tuning are both obtained by keeping the damping small, *i.e.*, by reducing to a minimum the rate of dissipation of energy during the oscillations of the diaphragm. The following factors contribute to the energy losses mentioned:—

(1) Deviations in the shape of the diaphragm from that of a perfect solid of revolution. Through the action of such defects of figure, there is a tendency for the energy originally contained in the fundamental mode to be dissipated in higher unsymmetrical overtones.

(2) Deviations from a thoretically rigid mounting and from geometrically perfect boundary conditions, giving rise to an unsymmetrical and varying tension in the annulus. As in (1), there is a tendency for the energy of the fundamental to be transferred to and dissipated in inharmonic overtones. In addition, there are energy losses at the boundary contact between the diaphragm rim and the mounting.

(3) Losses due to the elastic hysteresis of the material employed in the construction of the diaphragm.

(4) Energy dissipation at the microphone contact.

The damping due to factors (1) and (2) may be minimised by very precise workmanship. Special grinding tools should be constructed to give an almost optical precision to the figure of the diaphragm. In particular, special attention should be paid to the manner in which the diaphragm is fitted into position on the mounting. Not only should the surface of contact be fine-ground into position, but all traces of an oil-film should be removed. The latter has been observed to give rise to considerable damping through dissipation arising in a thin viscous film, subjected to alternating compressions while the diaphragm is vibrating. Anything in the nature of rubber packing, cemented or soldered contacts, should be avoided for the same reason.

The losses due to elastic hysteresis in the material composing the diaphragm can only be minimised by the choice of a suitable steel or alloy; the material should at the same time possess as high an elastic limit as possible to give a wide range of tuning. The selection of a suitable material should be made the subject of a special investigation by constructing diaphragms of various metals and measuring their free damping constant by some such instrument as Prof. Dayton C. Miller's "phonodeik."*

The elimination of losses at the microphone contact can only be achieved by carrying out a special investigation.

It will be noticed that contributions to the damping arising from factors (1), (2), and (4) are not affected by the process of tuning by gaseous pressure. There is a possibility that elastic hysteresis losses during vibration may increase with the tension in the annulus to the detriment of sharp tuning at the upper limit of the frequency-range.

---

* Described in 'The Science of Musical Sounds' (Macmillan and Co., New York).

A Method of Analysing Galvanometer Records.

By W. Hartree, M.A., and A. V. Hill, Sc.D., F.R.S.

(Received March 9, 1921.)

In recent papers,* we have described experiments in which photographic records were made of the deflection of a sensitive galvanometer registering the rise of temperature of a muscle (or other substance) resting on a thermopile. The system is a complicated one, involving the inertia, damping and control of the moving system of the galvanometer, and the heat conductivity and capacity both of the substance on the thermopile, and of the material of the thermopile itself; consequently the accurate determination of the course of the production of heat in the substance on the thermopile, from the photographic record of the galvanometer deflection, would appear to require the solution of complicated differential and integral equations, with a large number of constants to be determined by observation. To avoid this necessity we adopted a procedure in which a given amount of heat was liberated, practically instantaneously, by a powerful electric current passing for a short time through the substance on the thermopile, and the resulting deflection recorded photographically. This record, which we will call the "control curve," involves all the constants of the system, and may be used (according to the procedure described below), for the determination of the course of the heat-production in general.

The motion of the moving system of a sensitive galvanometer (assumed completely damped) is governed by linear differential equations with constant coefficients, as also is the flow of heat from substance to thermopile, and in the thermopile itself. Consequently if a deflection of the galvanometer following the course

$$y = \delta \mathrm{H} f(t), \qquad \text{(I)}$$

be given by heat $\delta\mathrm{H}$ liberated at time 0, a deflection

$$y' = \delta' \mathrm{H} f(t-\theta)$$

will be given by heat $\delta'\mathrm{H}$ liberated at time θ, and a deflection

$$y+y' = \delta \mathrm{H} f(t)+\delta' \mathrm{H} f(t-\theta)$$

by the combined effects of heat $\delta\mathrm{H}$ at time 0 and of heat $\delta'\mathrm{H}$ at time θ. In making the "control curve" we have determined experimentally the form of the function f in equation (1), and if we know the form of the function ϕ in the equation

---

* W. Hartree and A. V. Hill, 'Journal of Physiology,' vol. 54, p. 104 (1920); 'Phil. Trans.,' B, vol. 210, p. 153 (1920).

$$H = \phi(\theta),$$

where H is heat, and θ is time, we can at once compute the deflection at time t from the equation

$$y = \int_{\theta=0}^{\theta=t} \frac{d\phi}{d\theta} f(t-\theta)\, d\theta = \psi(t) \text{ (say).}$$

The converse problem—given the forms of the functions f and ψ, to determine the form of the function ϕ—is more difficult, but it can be carried out numerically in the following manner. From the "control curve" a "control table" is constructed giving the deflections at equal intervals (say of 0·1 second) corresponding to an instantaneous production (at time 0) of 1·0, 0·9, 0·8, 0·7, ... 0·2, 0·1, 0·05 units of heat, the unit being appropriately chosen. The observed curve of galvanometer deflection is then measured up, and the ordinates at intervals of 0·1 second tabulated. From inspection of the first two or three of these it is seen that they agree more or less with the numbers in the control table corresponding to heat (say) 0·7 units liberated at time 0; these numbers in the control table are then subtracted throughout from the tabulated ordinates of the observed curve, and a note made "0·7 units at time 0." The first few remainders are then seen to correspond to heat (say) 0·2 units at time (say) 0·2 second, and the numbers, therefore, for 0·2 units in the control table are shifted along to 0·2 second and subtracted throughout, leaving a second series of remainders. The first few of these remainders are then seen to correspond to heat (say) 0·1 at time (say) 0·3 second, and so on. Finally, therefore, it is found that by taking heat 0·7 units at 0 seconds, 0·2 at 0·2 second, 0·1 at 0·3 second, and so on, and subtracting, the remainders become zero or practically zero all along and the problem is solved.

If there be reason to suppose that the heat produced is a reasonably simple function of the time, the results of the first analysis may be plotted and smoothed, and a second attempt made. After about three attempts a final solution is usually arrived at. If the heat be a complicated function of the time, the analysis is apt to be difficult and unsatisfactory, but in most of the cases we have dealt with the method has proved reasonably accurate, if laborious. We have employed it in the determination of the heat produced by a muscle (only about 0·005 grm. calories per gramme) during the first two seconds, or during the first three minutes following stimulation.

The method requires very accurate registration of the galvanometer response—naturally, for it rests largely on an examination of quite small differences—and we doubt whether it could be employed successfully unless the records were made photographically; it is possible moreover only in cases where a sudden outburst of heat can be provided at a definite moment in order to produce the control curve—a condition difficult to satisfy when employing a non-conducting substance like rubber. In all cases, however, where such a "control" curve can be made, and where the motion of the galvanometer can be assumed to be governed by linear differential equations with constant coefficients, the method may be employed; it is not limited necessarily to the case of galvanometer deflections produced in response to changes of temperature in a thermopile.

On the Influence of Low Temperatures on the Magnetic Properties of Alloys of Iron with Nickel and Manganese.

By Prof. KAMERLINGH ONNES, For. Mem. R.S., Sir ROBERT A. HADFIELD, Bart., F.R.S. and Dr. H. R. WOLTJER.

(Received March 4, 1921.)

SECTION 1.—*Aim of the Research.*

Since its discovery by one of us in 1883, the material now universally known as manganese steel, that is the iron-manganese alloy containing about 12 to 14 per cent. manganese and 1·10 to 1·25 per cent. carbon has, on account of its many remarkable properties, been the subject of many researches.

2. The investigation of its magnetic behaviour in connection with thermal and mechanical treatment has taken a prominent place among them, as this method of investigation, perhaps more than any other, promised to give an insight into the constitution of steel alloys, and accordingly to show how their properties might be influenced in any desired direction. Manganese steel has not only been important in itself, but by reason of the light it has thrown upon the laws relating generally to steel alloys.

3. Manganese steel as cast or forged is non-magnetic and when this steel is treated in the ordinary way to obtain its peculiarly high toughness and tenacity, that is, after being heated to 1000° C. and quenched in water, it is only very slightly magnetic, namely, equivalent to less than one part in a thousand as compared with Swedish charcoal iron. It can, however, be made to reach a specific magnetism about two-thirds of that of pure iron by prolonged heating at 520° C. Cooling down to the temperature of liquid

air ($-182°$ C.) had practically no effect on this steel provided it was either in the non-magnetic or strongly magnetic condition, but the magnetism increased, if it was in an intermediate condition before immersion.

4. Now the non-magnetic condition of manganese steel has been considered as analogous to that of some nickel-iron alloys. As is well known, some of these latter alloys may be non-magnetic at ordinary temperature, but can be rendered magnetic at the same temperature, by cooling them to a sufficiently low temperature and allowing them to return to normal temperature. It was suggested therefore, that possibly the temperature to which manganese steel must be cooled down in order to become magnetic, might be still lower than that of liquid air. Although this view was probably not correct, as pointed out by one of us, it was thought to be interesting to cool the manganese steel down to the temperature of liquid hydrogen ($-253°$ C.) or even further to that of liquid helium ($-269°$ C.) to ascertain whether, starting with the material in its non-magnetic condition, it could be made magnetic at ordinary temperatures.

5. A preliminary experiment at liquid hydrogen temperature showed that this change, as expected by some, did not occur. The question, however, remained for other non-magnetic manganese-iron alloys, containing varying percentages of manganese, with carbon of fairly constant but small percentages; and also the material known as "Resista," that is, the iron alloy containing about 5 per cent. of manganese and 15 to 25 per cent. of nickel. We resolved, therefore, to extend the cooling experiments to such a series of manganese alloys and for purposes of comparison to include a series of nickel-iron alloys (these also having a low and fairly constant carbon percentage). The analyses of the various sets of alloys investigated and prepared by one of the authors are shown in Table I.

6. As regards the nickel-iron alloys, if an accurate quantitative investigation had been the aim of the research, it would have been better, of course, to work with binary alloys as pure as possible, but for a general survey this material seemed sufficiently promising. The specimens were prepared in duplicate at the Research Laboratory of the Hadfield Works, at Sheffield, one set being retained for immersion in liquid air there, and the other sent to the Cryogenic Laboratory at Leiden for immersion in liquid hydrogen. In the preparation of the specimens and in the experimental work at Hecla Works the authors have been assisted by the following Members of the Research Staff at the works mentioned, namely, Messrs. S. A. Main, B.Sc., T. G. Elliot, F.I.C. and W. J. Todd, to whom the authors tender their best thanks for the assistance rendered in these researches, which have required much time and patience to carry out. The size of the specimens was about

Table I.

Mark.	Analysis.			
	C.	Si.	Mn.	Ni.
	Manganese-Iron Alloys.			
2423	Less than 0·1 per cent.	—	1·10	
1379 B/2	0·08	—	3·50	
2425	Less than 0·1 per cent.	—	4·10	
2426	Less than 0·1 per cent.	—	5·50	
1379 C/1	0·12		6·67	
2427	Less than 0·1 per cent.	—	10·50	
3156/2	0·14	—	11·41	
2428	Less than 0·1 per cent.		12·90	
2429	Less than 0·1 per cent.		15·70	
2430	Less than 0·1 per cent.	—	19·80	
1379 F	0·07	0·13	22·84	
	Manganese 5 per cent., Nickel 15 per cent.			
1109 D	0·60	0·84	5·04	14·55
	Manganese 6 per cent., Nickel 25 per cent.			
1414 B	1·18	Est., 0·56	6·05	24·30
	Nickel-Iron Alloys.			
1287 A	0·19	0·31	0·79	0·27
B	0·14	0·20	0·70	0·51
C	0·13	0·23	0·72	0·95
D	0·14	0·21	0·72	1·88
E	0·19	0·20	0·65	3·78
F	0·18	0·31	0·65	5·89
G	0·17	0·28	0·68	7·78
H	0·16	0·20	0·86	9·43
H/1	0·24	—	0·68	10·65
I	0·18	0·22	0·93	11·24
J	0·23	0·24	0·98	15·68
K	0·19	0·27	0·93	19·80
L	0·16	0·30	1·00	24·67
M	0·14	0·38	0·86	29·39
N	0·14	0·28	1·49	43·00

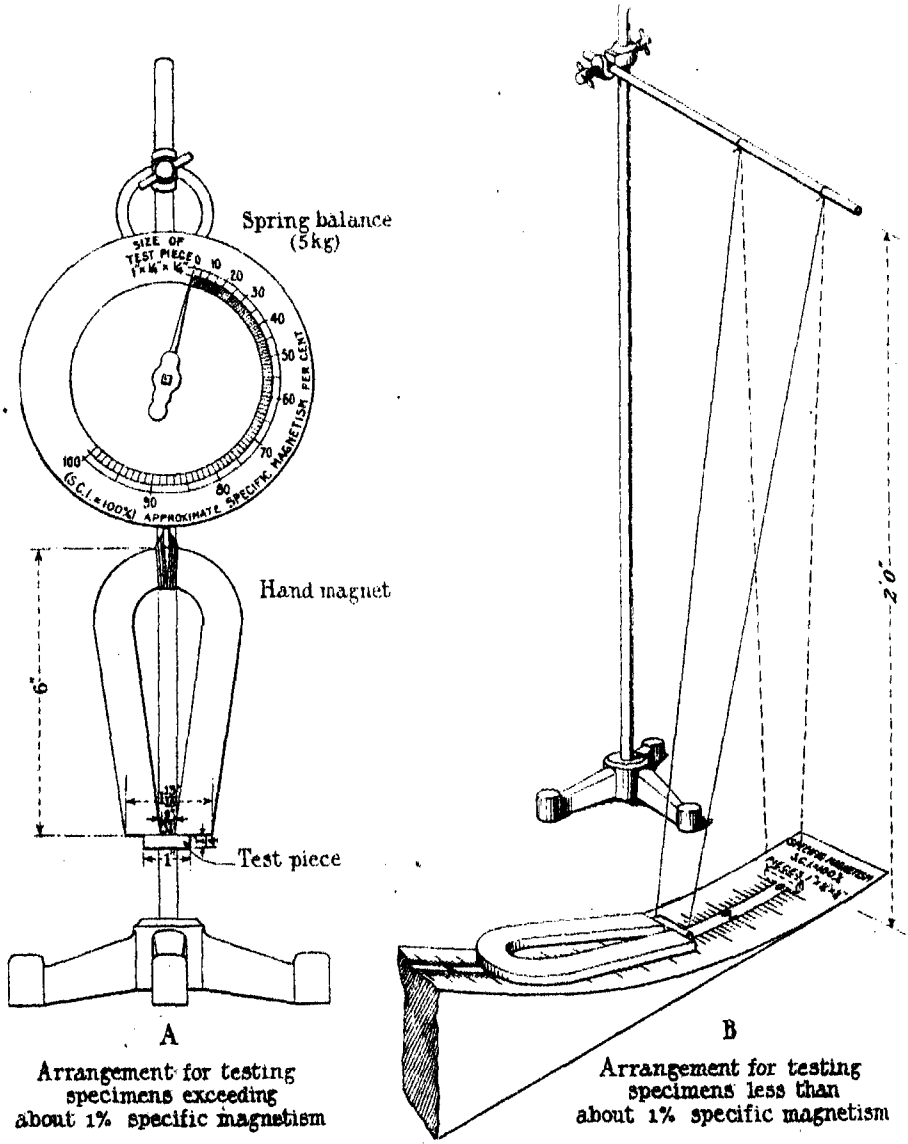
0·7 cm. square × 2·75 cm., and they were quenched from 1000° C. in water at atmospheric temperature.

SECTION 2.—*Method of Investigation.*

1. For the reason mentioned in Section 1, no complete magnetisation curves were recorded, nor was the saturation magnetisation measured by an accurate method as carried out by Hegg for pure iron-nickel alloys, but in the present case it was considered sufficient to measure the attraction in a non-homogeneous field by a simple method.

2. At the Hecla Works Laboratory the specimens were tested magnetically before and after immersion in liquid air, by the attractive force of a 6-inch

hand magnet placed with both poles in contact with the specimens, the force being measured either by a spring balance for the more magnetic, or by gravity for the less magnetic specimens (*i.e.*, less than about 1 per cent.


Spring balance
(5kg)
SIZE OF TEST PIECE 1"×¼"×¼"
0
10
20
30
40
50
60
70
80
90
100
(S.C.I.=100%) APPROXIMATE SPECIFIC MAGNETISM PER CENT
Hand magnet
6"
1"
Test piece
2'0"

A

Arrangement for testing specimens exceeding about 1% specific magnetism

B

Arrangement for testing specimens less than about 1% specific magnetism

Fig. 1.

specific magnetism, fig. 1). This method while necessarily approximate, has been found to give results agreeing fairly closely with those obtained from saturation tests by the Isthmus method. Tests on the actual Leiden specimens by this method also agreed closely with the results obtained at Leiden, except for a tendency to give rather lower values for the less magnetic specimens. The Brinell hardness was also measured before and after immersion.

3. As regards the measurements carried out at Leiden, it was desirable for a reason explained later (Section 5) to execute the measurements not only before and after the immersion in liquid hydrogen and helium, but also *during* the immersion. In order to determine the attractional forces on the different samples quickly, without losing too much hydrogen, it was necessary to construct a special apparatus, as it necessitates the removal of the specimen from the bath of liquid hydrogen, and its replacement by another, without opening the cryostat too much, or letting the air enter. We succeeded in realising these conditions by the following apparatus.

SECTION 3.—*Apparatus.*

1. Figs. 2 and 3 show the apparatus, which might be termed a kind of balance, the beam being rectangular in form. By placing the weights on the scale, S, the force which the field exerts on the magnetised bar in the box, B, is determined. The vacuum glass being filled with liquid hydrogen the experiment was carried out in the following manner:—The vacuum glass with cap was moved down along the glass tube, G, until the pointer, P, reached a scratch previously made on the glass. By now raising the box, B, from the shuttle, Sh, by turning the handle, H, which works on an eccentric, E, it comes just between the apertures C_1 and C_2; the tested bar can be pushed by means of a specially made rod from the box B through one of the apertures C, and replaced by another, as in fig. 4 (the new specimen is introduced from the right; immediately on the removal of the tested specimen, the cover D is screwed on).

2. The vacuum glass is then moved along the glass tube G until the pointer marks its original position. The rubber ring, R, usually closes sufficiently tight to prevent movement of the vacuum glass, and, moreover, the point of P rests in a small hole. Any movement of the glass is thus entirely prevented. All parts of the balance were made from non-magnetic material, except the steel pins on which the apparatus hangs. One of these pins rests in a hole, the other in a slot. When the balance is placed in the magnetic field, it is mounted on a cross-sledge for easy adjustment. When not in use, or when the vacuum glass is to be filled with liquid hydrogen, or when a new sample is to be inserted, the beam is removed from between the poles and placed on an auxiliary table, to spare the pins and to have more room for operating. The vapour rising from the liquid hydrogen boiling in G must have a vent in

order to obtain a quiet boiling; for this purpose the outlet at the top of G and the outlet of the vacuum glass are connected, by means of a rubber tube, to a T-shaped glass tube, through which the escaping vapours may be conducted, if necessary, to the gas-holder. When the measurements are to be made and the balance must be free, the T-piece is naturally disconnected from the gas-holder. The electromagnet was the great Weiss electromagnet of the Laboratory.

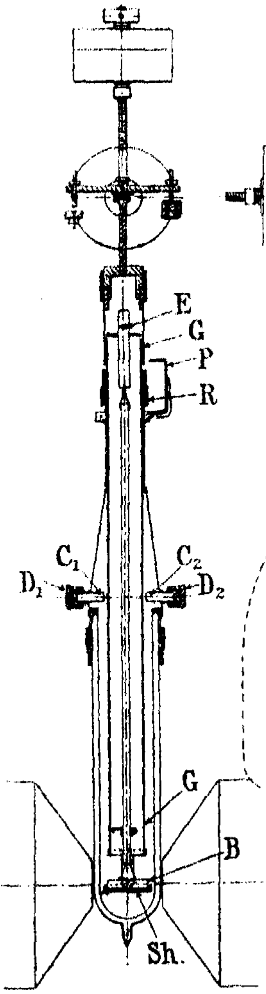


FIG. 2.

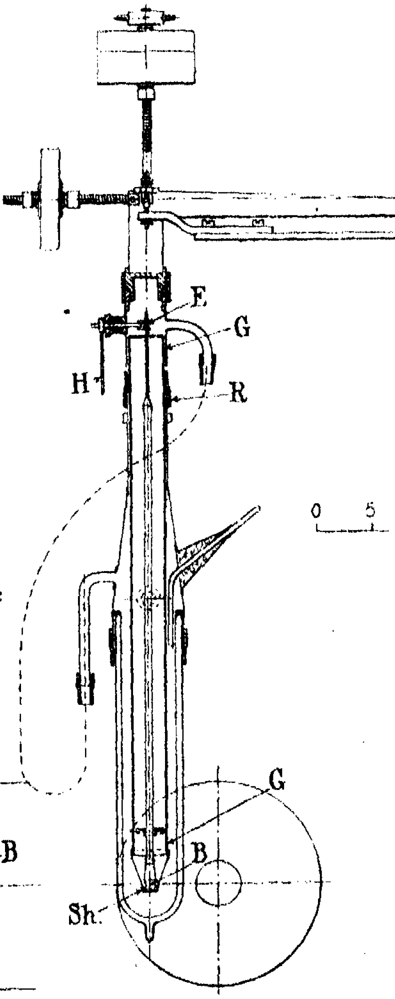


FIG. 3.

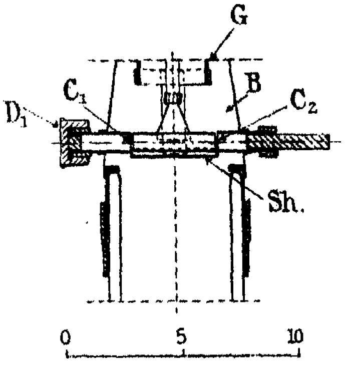


FIG. 4.

SECTION 4.—*Measurements.*

To execute the measurements, the cross-sledge with the bearings of the beam was arranged so as to make the edge of the balance parallel to the (horizontal) axis of the magnet and thus cut perpendicularly in the middle by the plane of symmetry passing between the poles. The steel pins being adjustable, it could be so arranged that the middle of the box, B, moved in this plane of symmetry, the bar contained in the box being always perpendicular to the plane. The counterweights on the beam were brought into position to cause the lever arm on which the scale S worked to be horizontal when the balance was at rest. The lower of the two screws S_1 and S_2, was raised until it only just touched the beam, the upper one was lowered to a few tenths of a millimetre above the beam, in order to prevent large oscillations.

2. The centre of the sample under investigation was in all cases at the same distance from the axis of the field, about 6 cm. The current through the electromagnet was then switched on and raised to 5 ampères. The field intensity in the centre (distance between the poles = 61 mm.) is about 4000 Gauss, the intensity of the external field in the centre of the bar about 2400, and the fall of intensity (dH/dx) about 280 C.G.S. The attraction exerted on the bar turned the beam to touch the screw S_1; by putting the weights in the scale it was made to rest on S_2. The weight required to bring the beam just into contact with S_2 was thus determined. The distance between the beam and S_1 being very small, this weight differed only very slightly from that needed to remove it from S_1.

3. A weight of P grammes in the scale corresponded with an attraction on the bar of 0·766 P grammes. P varied for the different samples from about 800 to 1 grm. and could be determined easily with an accuracy of 0·1 grm. The attraction exerted on the apparatus without the sample was quite negligible. P was determined for both directions of the current through the electromagnet, but the difference was always very slight: the effect of hysteresis of the samples was not appreciable.

4. For purposes of comparison, the attraction exerted on a sample of material known as Swedish charcoal iron (S.C.I.) of the same size as the other samples was also measured. For this measurement at the Hecla Works this very pure iron was used. A specially selected batch of about 2 tons of the material has been used by one of us in all his experiments during the last twenty-five years, and may be taken as characteristic of and representing the physical and other qualities of pure iron. We refer to it as "S.C.I." The following is its percentage analysis:—

	C.	Si.	S.	P.	Mn.	Fe (by diff.).	Total impurities.
"S.C.I." ...	0·045	0·07	0·005	0·004	tr.	99·876	0·124

5. The results obtained at Leiden originally referred to a rather less pure

iron, designated "L." The experiments at Leiden having been finished, a sample of it was analysed at the Hecla Works. It has the following composition:—

	C.	Si.	S.	P.	Mn.	Cr.	Fe (by diff.).	Total impurities.
"L" ...	0·05	0·03	0·013	0·068	Nil	0·09	99·749	0·251

6. Under the field-conditions indicated above, it was attracted 1 per cent. less than the pure iron from the Hecla Works. The original numbers obtained, therefore, which had reference to the "L" material, have been lowered by 1 per cent.

SECTION 5.—*Results and Discussion.*

1. *General Remarks.*—The results of immersion in liquid air at Hecla Works are shown in Table II and figs. 5 and 6; those of immersion in liquid hydrogen at Leiden in Table III and figs. 7 and 8. Beyond noting that the results at Hecla Works are in agreement in a satisfactory way with the results from liquid hydrogen at Leiden, it is not necessary to refer individually to them as regards the magnetism of the specimens. The irreversible ranges indicated both for the nickel and manganese steels also compare well with those shown in the Leiden experiments. As regards the Brinell hardness, the irreversible change both in the iron-nickel alloys and in the limited intermediate range of iron-manganese alloys is accompanied by a distinct increase in hardness. Where there is no marked change in the specific magnetism the hardness also does not show any marked change.

2. *Theory of Low Critical Points for the Manganese-Iron Alloys.*—The experiments show that the *non-magnetic manganese alloys do not become magnetic either by immersion in liquid air or by immersion in liquid hydrogen; neither are they magnetic at the temperature of liquid hydrogen itself.* (Here and in the following discussion the importance of the measurements made at the temperature of liquid hydrogen appears.) The 11·4 and 12·9 per cent. manganese alloys may be in two different magnetic conditions at atmospheric temperature; the latter, moreover, is remarkable in being more strongly magnetic after immersion than during immersion in liquid hydrogen. Apparently for these alloys atmospheric temperature lies between the two critical temperatures, liquid air temperature being below the lower critical point.

2. If the non-magnetic condition of the higher manganese-iron alloys is to be explained by one or both of their critical temperatures being exceedingly low, there are two possibilities:—

(*a*) Atmospheric temperature is between the two critical temperatures.

(*b*) Atmospheric temperature is higher than the upper critical temperature.

Table II.—Effect of Immersion in Liquid Air (−182° C.) at Hecla Works.

Specimens previously quenched in water from 1000° C.
The specific magnetism figures have reference to Swedish charcoal iron taken as 100.
The specimens were immersed in liquid air for 12 minutes.

Mark.		Before immersion in liquid air.		After immersion in liquid air.	
		Specific magnetism.	Brinell* hardness.	Specific magnetism.	Brinell* magnetism.
	Low Carbon Manganese-Iron Alloys.				
	per cent. Mn.				
2423	1·10	96	166	95	142
1379 B/2	3·50	92	370	91	394
2425	4·10	90	370	90	370
2426	5·50	89	370	90	370
1379 C/1	6·67	86	288	82	280
2427	10·50	50	394	60	480
3156/2	11·41	35	370	58	486
2428	12·90	8	274	20	318
2429	15·70	1	190	1	210
3430	19·80	0·42	199	0·42	214
1379 F	22·84	0·40	202	0·45	208
	Manganese 5 per cent., Nickel 15 per cent.				
1109 D		0·10	142	0·20	138
	Manganese 6 per cent., Nickel 25 per cent.				
1414 B		0·30	128	0·30	128
	Low Carbon Nickel-Iron Alloys.				
	per cent. Ni.				
1287 A	0·27	93	156	97	155
B	0·51	93	162	96	176
C	0·90	90	166	95	156
D	1·88	87	342	93	336
E	3·73	90	364	92	342
F	5·89	86	342	85	370
G	7·78	88	370	86	364
H	9·43	87	418	84	400
H/1	10·65	87	432	88	408
I	11·24	80	436	86	418
J	15·63	76	476	75	476
K	19·80	74	400	78	394
L	24·67	49	258	78	370
M	29·39	3	107	69	332
N	43·00	88	129	90	125

* 10 mm. ball, 1000 kgrm. load.

As regards (*a*), if the first hypothesis is true, our experiments show that the lower critical temperature must be still lower than the boiling point of liquid hydrogen. In order to test this idea, the non-magnetic manganese alloys were immersed *in liquid helium* (−269° *C.*) *for some two and a half hours, but, when they were again at atmospheric temperature, none of them*

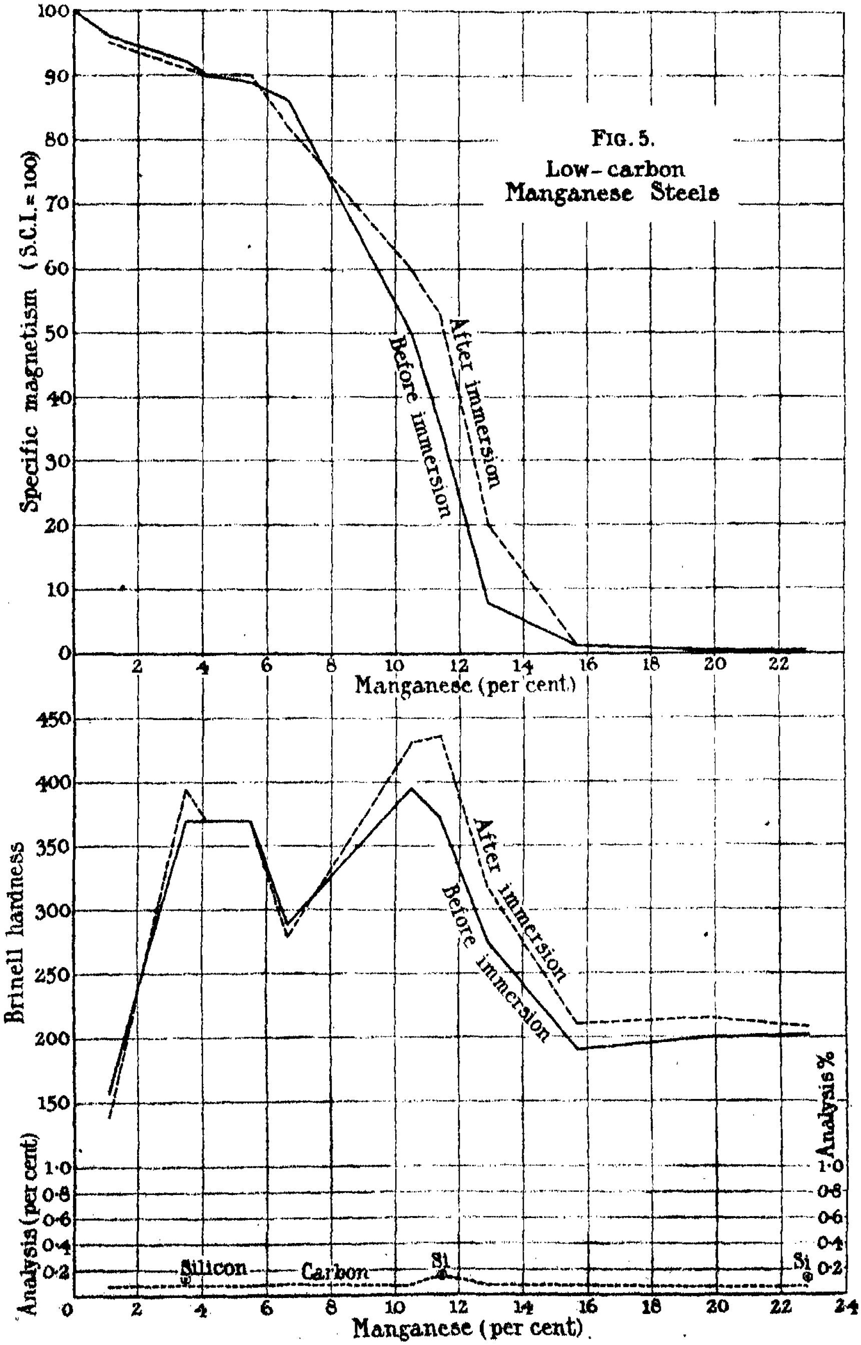


FIG. 5.
Low-carbon
Manganese Steels
Specific magnetism (S.C.I. = 100)
100
90
80
70
60
50
40
30
20
10
0
2
4
6
8
10
12
14
16
18
20
22
Manganese (per cent)
After immersion
Before immersion
Brinell hardness
450
400
350
300
250
200
150
After immersion
Before immersion
Analysis (per cent)
1·0
0·8
0·6
0·4
0·2
0
Analysis %
Silicon
Carbon
Si
Si
24
Manganese (per cent)

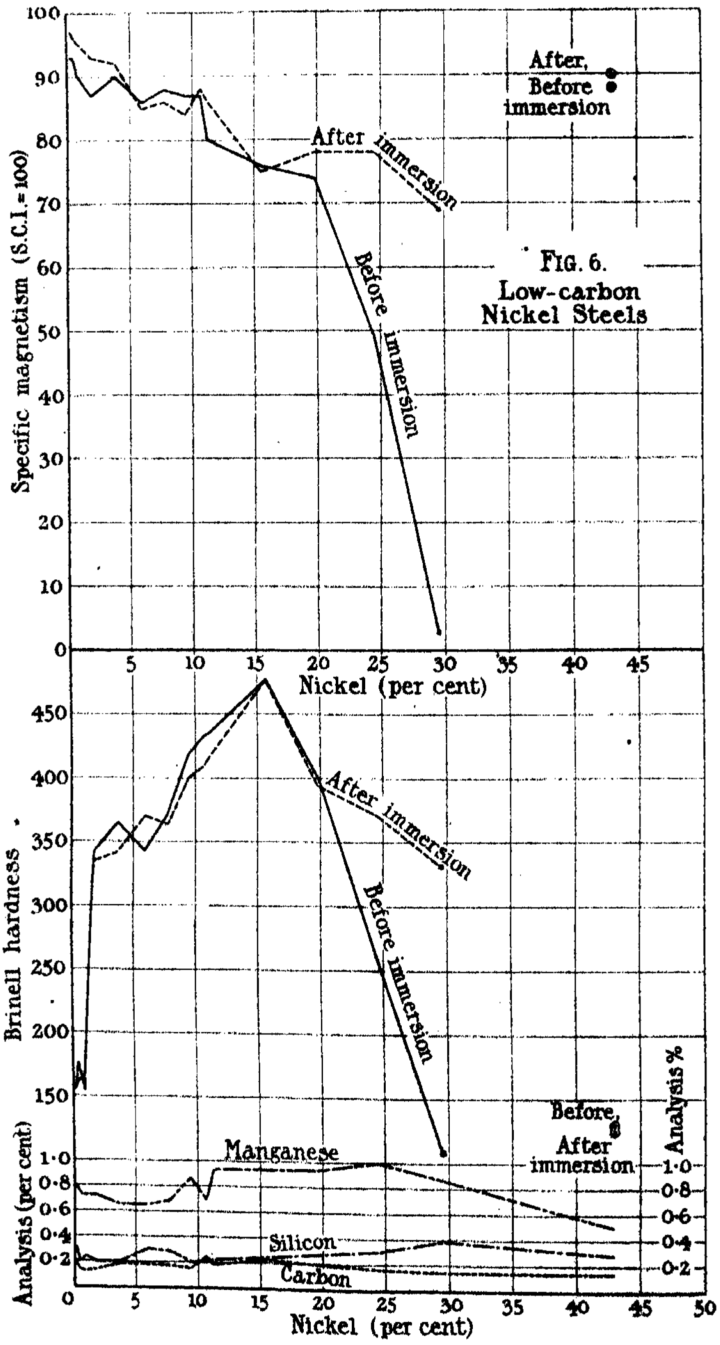


FIG. 6.
Low-carbon Nickel Steels

Table III.—Effect of Immersion in Liquid Hydrogen (−253° C.) at Leiden.

Specimens previously quenched in water from 1000° C.

The specific magnetism figures have reference to Swedish charcoal iron taken as 100.

Mark.		Specific magnetism.		
		Before immersion in liquid hydrogen.	During immersion in liquid hydrogen.	After immersion in liquid hydrogen.
		Low Carbon Manganese-Iron Alloys.		
	per cent. Mn.			
2423	1·10	97·0	99·6	97·4
1379 B/2	3·50	93·0	96·4	93·5
2425	4·10	94·2	96·7	94·3
2426	5·50	92·7	95·8	92·7
1379 C/1	6·67	91·1	94·2	91·5
2427	10·50	51·9	55·4	53·8
3156/2	11·41	38·5	46·0	46·2
2428	12·90	9·8	15·7	17·7
2429	15·70	2·9	3·1	3·1
2430	19·80	0·81	0·89	0·80
1379 F	22·84	0·53	0·59	0·53
		Manganese 5 per cent., Nickel 15 per cent.		
1109 D		0·77	1·74	0·96
		Manganese 6 per cent., Nickel 25 per cent.		
1414 B		1·49	8·78	1·57
		Low Carbon Nickel-Iron Alloys.		
	per cent. Ni.			
1287 A	0·27	96·2	98·0	96·4
B	0·51	97·5	99·4	97·8
C	0·90	97·5	99·9	97·9
D	1·88	94·4	97·1	94·7
E	3·78	95·0	97·9	95·8
F	5·89	94·8	98·3	95·1
G	7·78	94·8	98·0	95·2
H	9·43	96·3	99·2	96·4
H/1	10·65	90·3	93·8	90·6
I	11·24	93·5	96·8	93·9
J	15·63	84·6	88·7	85·9
K	19·80	78·3	86·6	83·0
L	24·67	48·1	83·6	79·8
M	29·39	8·9	72·1*	64·5
N	43·00	71·0	81·0	72·0

* *During a second immersion the specific magnetism was 76·2.*

showed any appreciable change in magnetic condition (Table IV). The attraction during the immersion in liquid helium was not measured on account of the extraordinary difficulties connected with all work at helium temperatures. It might be asked also if it was not necessary to keep the samples even longer in the liquid helium to obtain the magnetic transformation, since all transformations occur very slowly at low temperatures. It must be remarked, however, that the magnetic transformations seem to occur rather quickly, 12 minutes in liquid air being sufficient for many transformations (see Table II). It appears, therefore, that on hypothesis (*a*) the lower critical temperature for 15, 19·8, and 22 per cent. manganese must be lower than the boiling point of liquid helium (−269° C.). As that temperature, however, for 12·9 per cent. manganese lies above the boiling point of liquid air, there would be an enormous difference between the critical temperatures of two neighbouring alloys.

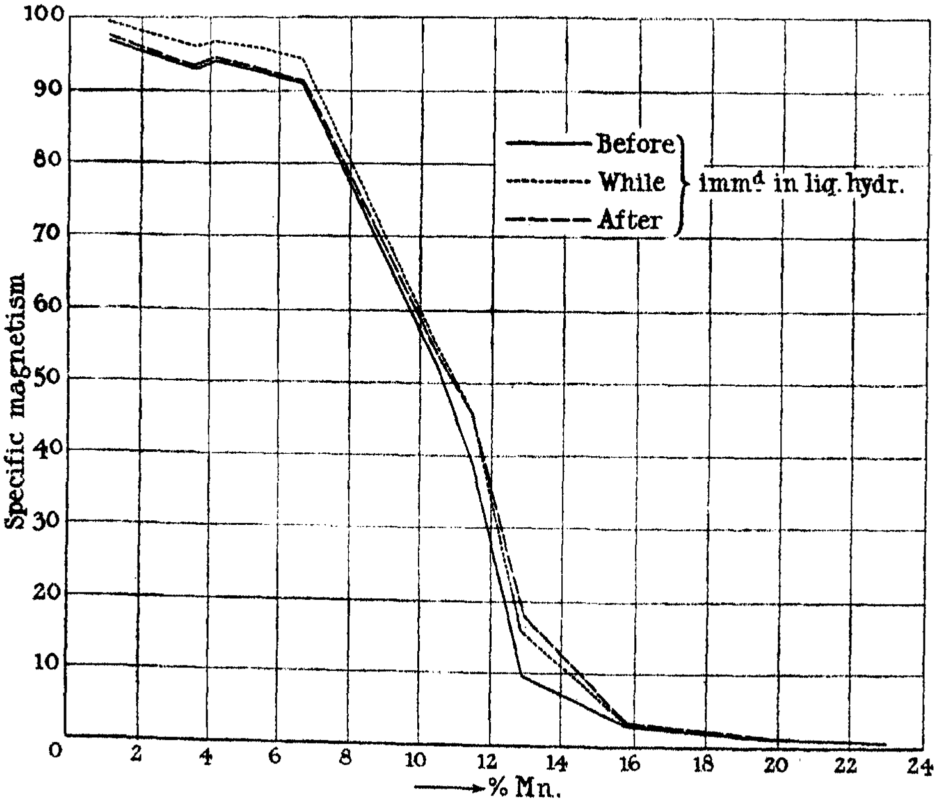


Fig. 7.—Specific magnetism iron-manganese alloys.

Regarding hypothesis (*b*), in this case we must conclude from our measurements in liquid hydrogen that the lower critical temperature must be below the boiling point of liquid hydrogen, since the alloys appear still to be non-magnetic at that temperature. The immersion in liquid helium being unaccompanied by a measurement in liquid helium itself does not in this case prove anything.

Though hypothesis (*b*) fits the results of this research quite well (see fig. 8) still it seems as little acceptable as hypothesis (*a*), if we take into account some unpublished experiments of one of us, and of the observations of

Gumlich, who determined the magnetic critical points themselves by a direct method, and whose results may be seen from fig. 9. Moreover, there appears to be an explanation, as given below, which has many advantages over the theory of the low critical points.

3. *Suggestion of two Manganese-iron Compounds.*—The form of magnetism curve for the iron-manganese alloys in figs. 5 and 7 is very suggestive. Considering only the curves representing the condition stable at low temperature, that is, after immersion in liquid air or liquid hydrogen, from zero manganese to about 7 per cent., the magnetism falls in a practically linear manner. From about 7 per cent. manganese to about 16 per cent. manganese, the fall in magnetism is again approximately linear, but at a very much more rapid rate.

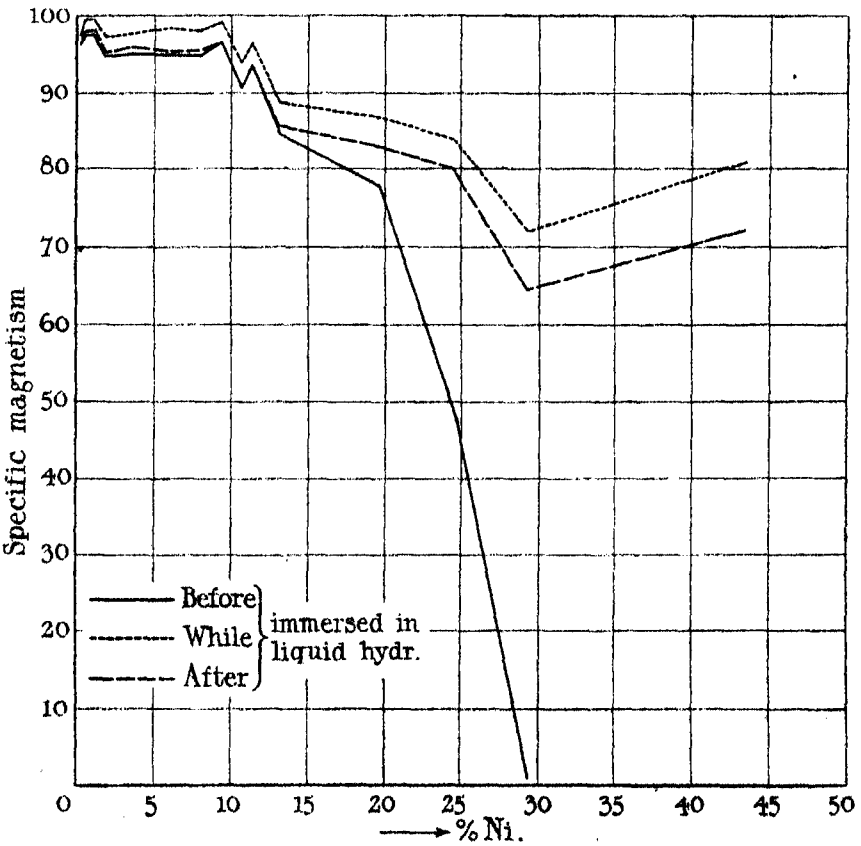


Fig. 8.—Specific magnetism iron-nickel alloys.

This strongly suggests definite chemical compounds of iron and manganese, containing respectively about 7 and 15 per cent. manganese. From this

Table IV.—Effect of Immersion in Liquid Helium (−269° C.) at Leiden.

Specimens previously quenched in water from 1000° C., except where otherwise stated.
The specific magnetism figures have reference to Swedish charcoal iron taken as 100.
The specimens were immersed for about 2½ hours.

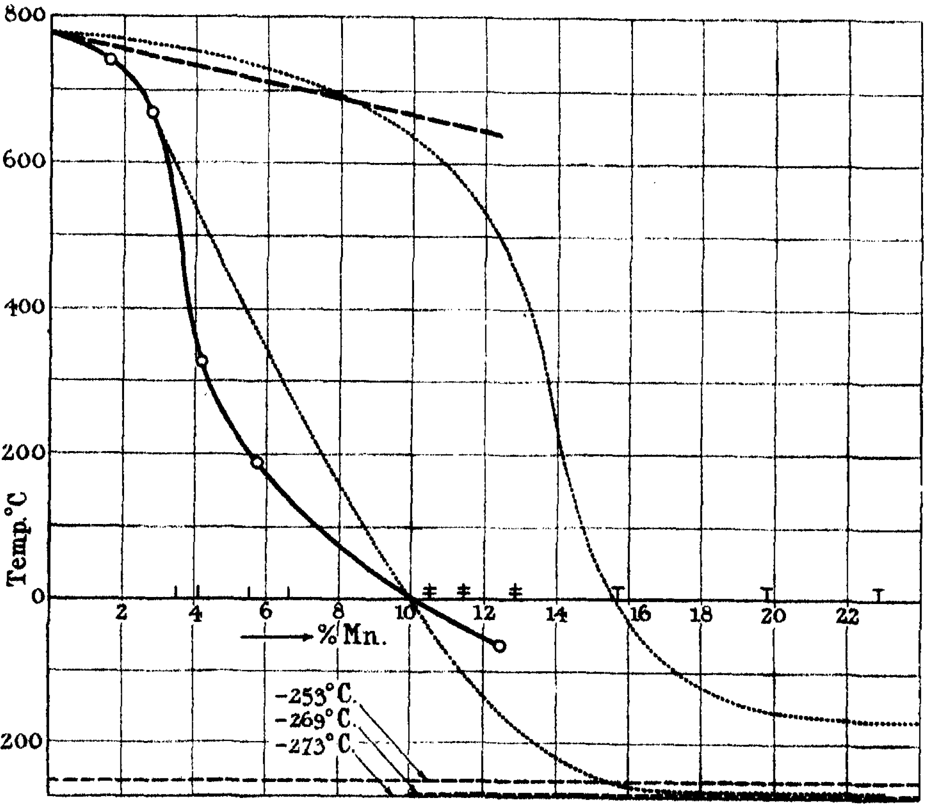
Mark.	Composition.	Specific magnetism.	
		Before immersion.	After immersion.
	Per cent. Mn.		
2428	12·9	17·7	17·7
2429	15·7	3·3	3·3
2430	19·8	0·70	0·72
1379 F	22·8	0·50	0·50
1109 D	0·60 per cent. C; 5·0 per cent. Mn; 15 per cent. Ni	0·99	0·98
1414 B	1·18 per cent. C; 6·0 per cent. Mn; 25 per cent. Ni	1·51	1·51

Mark.	Heat treatment.	Specific magnetism.	
		Before immersion.	After immersion.
Manganese steel B ...	As forged ...	0·13	0·12
Manganese steel D 1·20 per cent. C. 12·96 per cent. Mn	500° C. for 60 hours, then re-forged and water-toughened.	0·13	0·16

point of view, the alloys containing from 0 to 7 per cent. manganese would consist of a mixture of pure iron, and the less manganiferous compound of the two; between 7 and 16 per cent. manganese there would be a mixture of the two iron-manganese alloys; and, above 16 per cent., of a mixture of the more manganiferous compound with pure manganese. From such an assumption, the compounds would have a specific magnetism of about 85 per cent., and practically zero, respectively.

The presence of such intermediate compounds might further go to explain the behaviour of the alloys as regards the presence or absence of critical points on heating and cooling. Above 16 per cent. manganese, the non-magnetic (or at most only slightly magnetic) iron-manganese compound, is mixed with the non-magnetic element manganese. Therefore no critical magnetic change points are to be expected in the range above this percentage. Approaching 16 per cent. manganese from the lower percentages, the proportion of the less manganiferous and magnetic iron-manganese alloy (about 7 per cent. manganese), admixed with the non-magnetic iron-manganese alloy (about 16 per cent. manganese), becomes less and less; the main proportion of the material consists of the non-magnetic constituent, which has no critical points; consequently, the original magnetic change points in the material are evanescent.

A reference to the curve of Brinell hardnesses in fig. 5 shows minima of hardness at about 7 and 16 per cent., passing through maxima at points between 0 and 7 per cent. and 7 and 16 per cent. This again would be in conformity with the existence of the two compounds of iron and manganese referred to, following the usual behaviour of simple mixtures.



═══ } Magnetic transformation-temperatures according to Gumlich.
·········· " " temperatures (not observed) agreeing with hypothesis C.

Fig. 9.

It is well known that, as a general rule, when two elements, and element and a compound, or two compounds, intermix in all proportions, the hardness of the mixture passes through a maximum as the percentage of one of the constituents is increased gradually from 0 to 100 per cent.

The data are insufficient to justify definite conclusions as to the existence of these compounds or their precise composition; the specimens are insufficient in number, especially in the critical regions, for example between 6·7 and 10·5 per cent., and, again, between 12·9 and 19·8 per cent. The curves, therefore, which have been drawn by joining up the points of observation, only roughly approximate to the complete curves.

The compound represented by Fe_5Mn would contain about 16·4 per cent. manganese; that represented by $Fe_{12}Mn$, 7·6 per cent. manganese. Fe_6Mn would contain 14·1 per cent. manganese, and would appear to be a more likely compound from the point of view of valency, but with a higher percentage than this, namely, specimen 2429, with 15·7 per cent. manganese, there is still a very small, though definite, amount of magnetism in the material.

It will be understood, therefore, that the suggestion as regards these compounds is put forward in quite a tentative manner, as likely to provide a more satisfactory alternative to the theory that the non-magnetic character of the more manganiferous iron-manganese alloys is due to the progressive lowering of the critical magnetic change point of the iron, even below the lowest temperatures practically attainable.

It is introducing no new principle to suggest that iron, by combination with another element, may entirely lose its ferro-magnetic character. On the other hand, the tendency to associate the magnetic character simply with changes of micrographic structure connected with the critical points disregards any physical differences which there may be between materials having similar micrographic structure.

It may be added that our results agree in many respects with those of Gumlich, who found for the saturation magnetism of manganese-iron alloys, after quenching from 800° C., two straight lines. The saturation magnetism in iron decreases with increasing percentages of manganese, linearly at first, that is up to about 7·5 per cent. manganese, slowly then from 7·5 per cent., rapidly down to zero at 15 per cent. manganese. Gumlich finds, however, that the decrease of magnetism up to about 8 per cent. manganese is proportional to the percentage, and that the manganese acts as a diluent only in these alloys. As regards the higher alloys, he mentions also the suggestion by Hilpert of a non-ferromagnetic compound, between iron and manganese, but his theory seems to differ from that just given. Gumlich observed discontinuities also in the densities and in the electrical resistance, though not all exactly at the same percentages of manganese.

Micrographical examination does not give any indication of a special constituent at the critical percentages indicated by the physical properties. This, however, is consistent with the compounds having a similar crystalline character to the component elements, iron and manganese. Such cases occur with some other elements which form compounds.

The abrupt bend in the curves at 7 per cent. manganese must indicate something special for this composition, as its critical temperatures are well above atmospheric temperature, and this abrupt change cannot be accounted for by the same explanation as applies in the case of nickel-iron alloys.

As regards the composition of such compounds, supposing they exist, while the compounds already established in the case of other metals usually have a simple atomic proportion, there do exist several compounds in which one of the elements greatly predominates. Examples of these are given by one of the latest authorities (Giua), so that the compounds of iron and manganese suggested would not be unique in this respect. The same authority also points out that the formulæ for inter-metallic compounds are not necessarily determined by the ordinary valencies of the separate elements, so that *prima facie* there is nothing to preclude even compounds of the type Fe_6Mn, though, as mentioned, nothing really definite is suggested as to the composition of these compounds. If the suggestion lead investigators to study the matter further, or if it lead to a more satisfactory theory, it would appear to be fully justified.

4. *Effect of Carbon.*—In view of the important effect of carbon in manganese steel, it was thought possible that the presence of even 0·10 per cent. of carbon in the iron-manganese alloys forming the subject of this research might have an appreciable influence on the results obtained. It has been previously shown by one of us that by annealing manganese steel at about 500° C. for several hours the carbon is deposited as a carbide, forming an acicular structure.

Arnold has also isolated the carbides from annealed manganese steel, and where the manganese percentage is about 12 per cent. he finds the analysis of these carbides to correspond to the form $3Fe_3C.Mn_3C$. There is a probability that in manganese steel containing about 15 per cent. of manganese or over the double carbide has the formula $2Fe_3C.Mn_3C$. In either case, 0·10 per cent. of carbon corresponds to over 1 per cent. of the double carbide.

This method of precipitating the carbon as carbides suggested a means of rendering the alloys more pure, at any rate in respect of carbon; the carbide having been precipitated (fig. 10) leaves a matrix much freer from carbon. As, further, the carbides separated chemically from alloys containing above 10 per cent. manganese were found to be non-magnetic, their presence would not interfere with the magnetic measurements except to act as a diluent to the relative amount by which they were present. Subject to this slight dilution therefore, the magnetic qualities, after this treatment, of the alloys above about 10 per cent. manganese represent the qualities of fairly pure iron-manganese alloys. The question, therefore, as to what extent the

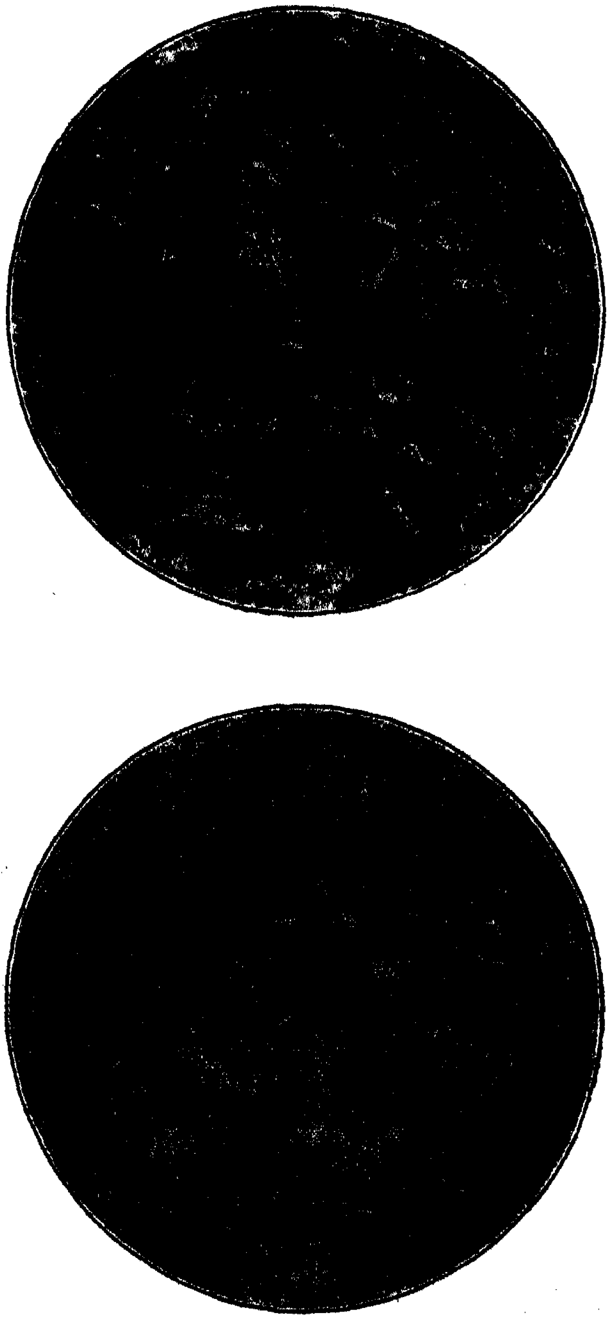


Fig. 10. ×600

irreversible range noticed in the iron-manganese alloys is influenced by, or due to, the amount of carbon present could be further investigated.

No inferences can be drawn, however, as to their magnetic qualities below about 10 per cent. manganese, as the deposited carbides are in these cases more or less magnetic, and mask the qualities of the matrix. The hardness over the whole range of the alloys is considerably influenced by the reinforcement of these hard carbides, so that also no deductions can be made as to the variation of the hardness of the matrix with variation in manganese percentage. Hardness and magnetic tests were, however, made on the whole series as a matter of interest.

Specimens $\frac{5}{8}'' \times \frac{3}{8}'' \times 2\frac{3}{4}''$ which had previously been quenched in water from 1000° C. were heated to 500° C. for sixty hours and cooled slowly. Magnetic and Brinell hardness tests were then made at atmospheric temperature, and the specimens immersed in liquid air for twelve minutes, the tests being repeated on return to atmospheric temperature.

The results are shown in Table V and fig. 11. The important result is now obtained that only three of the specimens now show a higher specific magnetism after immersion than before.

Table V.—Effect on Immersion in Liquid Air (−182° C.) on Specimens quenched from 1000° C. in water, then re-heated to 500° C. for 60 hours and cooled slowly.

The specific magnetism figures have reference to Swedish charcoal iron taken as 100.
Size of specimens: $\frac{5}{8}'' \times \frac{3}{8}'' \times 2\frac{3}{4}''$.
The specimens were immersed in liquid air for 12 minutes.

Mark.	Manganese (per cent.).	Before immersion in liquid air.		After immersion in liquid air.	
		Specific magnetism.	Brinell* hardness.	Specific magnetism.	Brinell* hardness.
	Low Carbon Manganese-Iron Alloys.				
2423	1·10	97	126	96	125
1879 B/2	3·50	88	268	92	263
2425	4·10	86	252	88	248
2426	5·50	86	258	86	266
1879 C/1	6·67	88	286	79	282
2427	10·50	40	350	40	364
3156/2	11·41	32	342	34	364
2428	12·90	7	286	9	306
2429	15·70	0·7	185	0·80	185
2430	19·80	0·28	231	0·20	260
1879 F	22·84	0·38	223	0·25	245

* 10 mm. ball, 1000 kgrm. load.

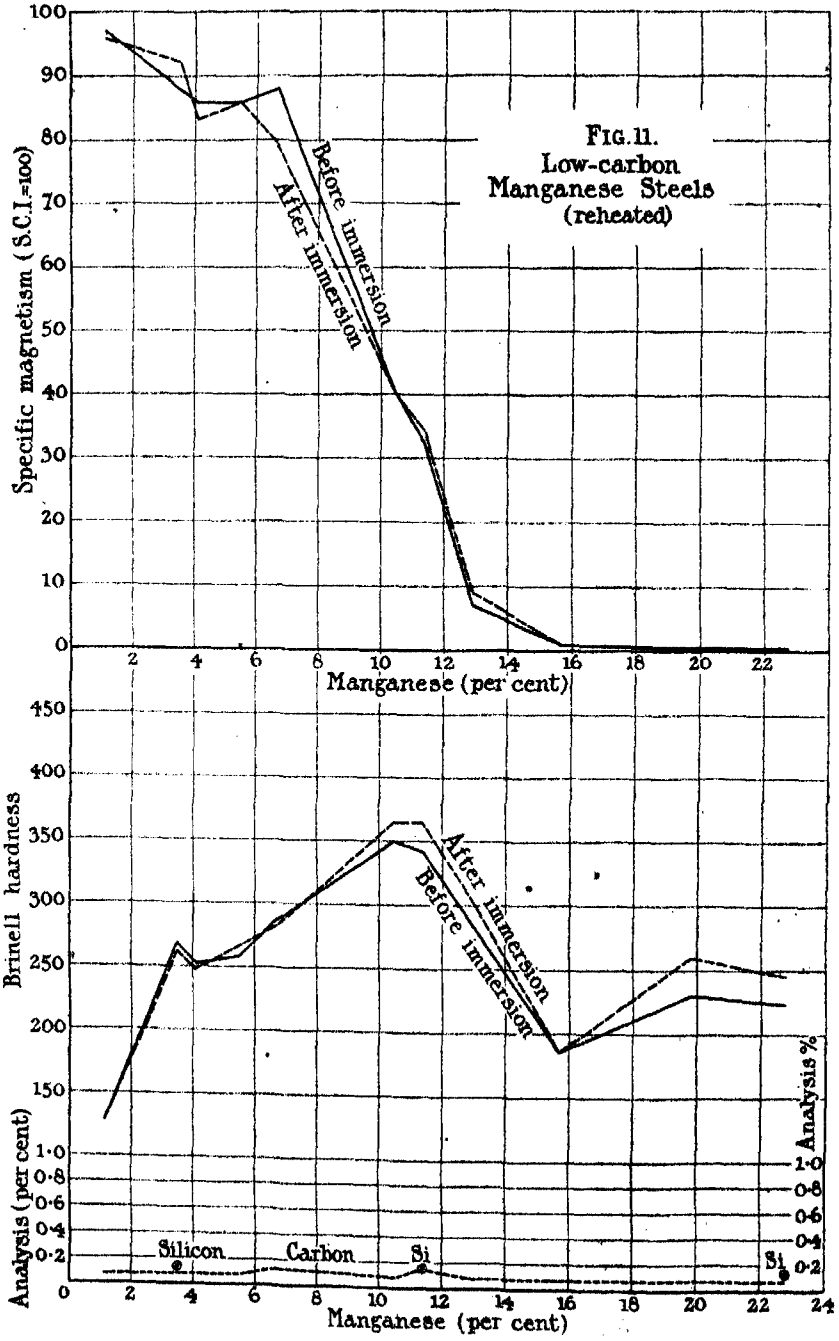


FIG. 11.
Low-carbon
Manganese Steels
(reheated)

One of these, No. 1379B/2 contains only 3·5 per cent. manganese, and is in the range affected by the presence of magnetic carbides; the other two, No. 3156/2 and 2428, are those which previously indicated an irreversible range. It will be noted that, for these, the increase of magnetism after immersion in liquid air is now very trifling. Consequently, with the elimination of carbon from the alloys, that is, leaving practically pure alloys of iron and manganese, it would appear that there is no irreversible range and that the magnetic properties at atmospheric temperature are a fairly fixed quantity for each alloy, without reference to their previous history.

As regards the Brinell hardness, as mentioned, no inferences can be drawn as to the probable relation of hardness to manganese percentage in the pure alloys; it is interesting, however, to note that the precipitation of the carbide has had the effect of making the changes due to immersion in liquid air less marked. That the minimum of hardness at about 7 per cent., noted in fig. 5, and taken as additional evidence in favour of the presence of iron manganese compounds, does not now appear, is probably due to the considerable influence of the deposited carbides in determining the Brinell number.

5. *The other Alloys.*—The *manganese-nickel steels* (known as "Resista" steels) are more strongly magnetic at hydrogen temperature than at atmospheric temperature, as generally happens in a small degree with ferro-magnetic, and in a higher degree with para-magnetic substances. Their magnetism at atmospheric temperature has not appreciably changed by immersion in liquid hydrogen nor by immersion in liquid helium, to which latter experience they were subjected, together with the pure manganese-iron alloys (Table III).

Two specimens of true manganese steel, *i.e.*, the ternary compound differing in its content of 1·25 per cent. carbon from the alloy 2428 included in the series of manganese-iron alloys (Table III) in its non-magnetic condition, have been immersed also in liquid helium, but without appreciable effect on their magnetism at atmospheric temperature (Table IV).

With regard to the *iron-nickel alloys*, these show the well-known irreversibility that might be expected, and are in this respect in accordance with the work of other authors.

Summary.

1. A series of iron-manganese and iron-nickel alloys, with a range of percentages of manganese and nickel respectively, has been tested in order to investigate the influence of cooling to very low temperatures (liquid hydrogen and liquid helium) on their magnetic properties, especially to

ascertain whether the iron-manganese alloys, which are non-magnetic at atmospheric temperature, become magnetic by so doing.

2. With a view to the desirability of testing not only after, but also during immersion in liquid hydrogen, an apparatus has been constructed, which makes it possible to test the samples quickly one after another at a temperature of 20° K., with as little loss of hydrogen as possible.

3. The iron-manganese alloys, containing the higher percentages of manganese, cannot be made magnetic at atmospheric temperature by cooling to the boiling point of liquid hydrogen or liquid helium.

4. The existence of one magnetic and one non-magnetic, or at most slightly magnetic, manganese-iron compound is shown to be probable, and the non-magnetic properties of the higher manganese-iron alloys may be explained by their means.

The Hardness of Solid Solutions.

By Walter Rosenhain, B.A., D.Sc., F.R.S. (of the National Physical Laboratory).

(Received March 22, 1921.)

It is a well-known fact of metallurgy that the addition of one metal to another produces an increase of strength and hardness. In some alloys, this change of properties is accompanied by the formation of a new micro-constituent or phase, which is itself harder, and also, as a rule, more brittle than either of the constituent metals. In a very large and important group of alloys, however, the addition of the second metal, up to certain limits of concentration, does not lead to the formation of a second phase or constituent. Alloys of this type, when they have attained an equilibrium condition, consist of an aggregate of polyhedral crystals, homogeneous in composition so far as their micro-chemical behaviour indicates, and in most respects entirely similar to the constituent crystals of the pure metal, which forms the basis of the alloy. A typical example of this kind is furnished by the alloys of copper with zinc, containing up to about 30 per cent. of zinc. Alloys of this type are generally described as "solid solutions," on the ground that the state of intimate mixture which exists in the liquid (molten) solution of the two metals in one another is preserved in these alloys after solidification. In Continental language, such crystals are more frequently termed "mixed crystals" ("Mischkrystalle"), but the present author prefers to avoid this term on account of a possibly misleading interpretation.

A further fact in connection with metallic solid solutions is also well known, but appears to require explanation. A solid solution alloy is always harder and stronger than the pure metal of which it mainly consists, and frequently this difference in physical properties is very marked. Thus, annealed pure copper has a Brinell hardness number of 36, and a tensile strength of about 13 tons per square inch. An alloy of copper containing 30 per cent. of zinc in solid solution, on the other hand, in the corresponding annealed condition, has a Brinell hardness number of 57, and a tensile strength of 20 tons per square inch. It is the purpose of the present paper to suggest an explanation for this hardening and stiffening effect of the added metal in solid-solution alloys, and to show that this explanation leads to an inference which is in striking accord with well-known facts.

The explanation of the properties of solid-solution alloys which is put forward in the present paper is based upon a conception of the manner in which two metals may crystallise together in the form of solid-solution crystals. As a result of the X-ray analysis of crystal structure, mainly carried out by Sir William H. Bragg and W. L. Bragg,* and by A. W. Hull,† it is now known that, in a metallic crystal, the atoms of the metal are arranged in certain definite ways on a space-lattice, and that this arrangement may differ in detail from one metal to another, not only in regard to the nature of the lattice (cubical or otherwise) and the occupied points, but also in regard to the normal inter-atomic distance or spacing of the lattice. On the view now put forward, the crystals of a solid solution of metal B in metal A are built up on the same space-lattice as crystals of pure A, the sole difference being that a certain number of individual atoms of A are replaced by atoms of B. It is obvious that the extent to which such an arrangement will be possible must depend upon the relation which exists between the atoms of the two metals. If both tend to take up a similar space-lattice arrangement, and if their normal inter-atomic distances do not differ very much, a single crystal may be built up, upon the normal space-lattice arrangement of A, using atoms of A and B indiscriminately, up to a considerable proportion of B. If the similarity of the two kinds of atom is sufficiently great, then such an indiscriminate construction might easily be conceived as being possible in all proportions between the two kinds of atom. In that case, we should expect to find that the two metals could form homogeneous solid solutions when alloyed in any proportion ranging from pure A to pure B. Such cases are, of course, well known among pairs of very similar metals, such as iron and nickel or silver and gold.

---

* W. H. and W. L. Bragg, 'X-rays and Crystal Structure,' p. 173.

† A. W. Hull, "A New Method of Chemical Analysis," 'J. Amer. Chem. Soc.,' 1919.

Where the similarity, in regard to crystalline habit and arrangement, of the two kinds of atom is less complete, it would be anticipated that, up to a certain concentration at either end of the series, the atoms would be built together indiscriminately upon a single space-lattice, but that, beyond certain limits of concentration, the two kinds of atoms would become segregated into two kinds of different crystals, thus giving rise to a duplex structure of two distinct kinds of crystals, each consisting of saturated solid solution of one metal in the other. The degree of dissimilarity of the two kinds of atom would determine the exact concentration at which this separation occurs, and we would thus expect to find in metallic alloy systems all gradations, ranging from complete mutual "solid" solubility down to almost complete insolubility in the solid state; it is, of course, well known that such gradations are found among known alloys.

From the present point of view, however, interest centres on the determining factor which governs the limiting solid solubility of metal B in metal A. If the view of the structure of a solid-solution crystal indicated above is correct, it follows that the actual arrangement of the atoms on the space-lattice typical of metal A cannot exist entirely undisturbed when some of the atoms of A are replaced by atoms of B. The forces acting on any atom of A, for instance, when one of its neighbours in the lattice is an atom of B, cannot possibly be the same as those acting on an atom of A entirely surrounded by atoms of A; there must thus be some want of symmetry in the forces acting on an atom so placed, with a result which is perhaps best expressed by saying that such an atom will be pulled slightly out of its proper place, *i.e.*, that the space-lattice will be slightly distorted at such points. Now, it is evident that such distortion must affect the internal energy of the whole system, and this effect must increase with the increase in the amount of distortion which accompanies the presence of a larger number of atoms of B. A point will therefore be reached when the energy-content of the system as a whole will be less if the atoms of B are separated and arranged on a space-lattice of their own. Such an arrangement implies the existence of boundaries between two different kinds of crystal and the accompanying storage of energy in such boundaries, and it is the balance between the energy-content of the two geometrically possible types of arrangement which must determine the limiting concentration of the solid solution in each case.

With a full knowledge of the forces at work between adjacent atoms it would be possible to calculate the limiting concentration of any solid solution from the character of the constituent atoms; in the present state of our knowledge, all that can be said is that the greater the extent to which the space-lattice is distorted in the neighbourhood of "dissolved" atoms (atoms of "B") the lower will be the limit of solid solubility of B in A. During the process of formation of such a solid solution, a great deal must depend upon the uniformity with which the atoms of B are distributed through the space-lattice. If there is a concentration of the dissolved atoms in any region, then there must be a tendency for rearrangement to occur with a view to arriving at some symmetrical arrangement which shall correspond to the minimum energy-content. If such a distribution is really symmetrical, then an interesting result must follow, viz., that in a solid solution which has been allowed to attain equilibrium conditions, the dissolved atoms will be distributed through the crystals in such a way as themselves to constitute some sort of space-lattice of their own. This is a conclusion which is capable of experimental verification by means of X-ray analysis; when the result is obtained, however, it will be important to realise that the existence of such a secondary space-lattice arrangement of the dissolved atoms (B) does not imply the existence of separate crystals of B. The further question of how far a solid solution which has attained such a state of symmetrical arrangement approaches to the character of a chemical compound need not be discussed here.

The most important conclusion which emerges from what has been said above is that the "solid solubility" of metal B in metal A will be, to a first approximation, inversely proportional to the amount of distortion of the normal space-lattice of A caused by the substitution of an atom of B for an atom of A in that lattice. The importance of this consideration lies in the fact that this amount of distortion of the space-lattice of A by the presence of B also governs certain of the physical properties of the resulting alloy, particularly the hardness of the material and its resistance to mechanical deformation generally.

It is now well recognised that the ductility of metals, *i.e.*, their power of undergoing plastic deformation without rupture, resides in their essentially crystalline character. Plastic deformation takes place within the crystals of metals by a process of slip, occurring on gliding or cleavage planes,* whereby the crystals accommodate themselves to the new shapes forced upon them by the deformation of the mass. It is precisely the power of the crystals to undergo such slip which governs the ductility of the metal, and their power of resisting slip which determines their power of resisting deformation. If this view is correct, then anything which tends to hinder the free occurrence of slip on the principal planes of a crystal will increase the hardness (or power of resisting deformation) and will lower the ductility of the material. The application of any appreciable amount of plastic deformation (or "cold work") is known to act in this manner, and does so, according to the views of Beilby,* by the partial destruction of the crystalline arrangement of the atoms within the crystals.

---

* Ewing and Rosenhain, 'Phil. Trans.,' A, vol. 249 (1899).

The view which it is desired to put forward here is that any even slight distortion of the space-lattice of the crystals of a metal by the presence of "dissolved" atoms of another metal, in the manner indicated in the earlier part of this paper, must serve as a hindrance to slip on the crystal planes. Strictly speaking, in fact, the crystal "planes" have ceased to exist in such a slightly-distorted crystal. The very facility for slip which renders metallic crystals soft and ductile depends upon the smooth regularity with which the atoms are arranged on their space-lattice, and if this smoothness is lessened by the slight distortion of the lattice, it is obvious that slip will be rendered more difficult, and this difficulty will increase, to a first approximation, proportionately to the amount of distortion existing in the space-lattice.

Combining this last inference with the one already drawn above in regard to the factor governing the limiting solid solubility of one metal in another, we arrive at the following conclusion: that since the amount of distortion which the introduction of a given "dissolved" atom produces in the space-lattice of the solvent metal governs both the limiting solubility of the dissolved metal and the degree of hardening produced in the solvent metal, we should expect to find that *the hardening effect of one metal upon another in the form of a solid solution* should, to a first approximation be *inversely proportional to its solid solubility*.

This very general conclusion, which—so far as the author is aware—has not been previously stated, appears to be in good general agreement with experimental fact. There are, in the first place, a certain number of binary alloy systems which consist of unbroken series of solid solutions. In some of these, of which gold-silver and iron-nickel are good examples, the shape of the liquidus curve suggests that the two kinds of atoms are very similar, since the alloys solidify in a manner closely approximating to that of a pure metal. In these cases there is very little hardening effect of one metal upon the other. The alloys of copper and nickel also form a series of unbroken solid solutions, but their freezing range is distinctly larger, so that the alloys depart further from the behaviour of pure metals. Here, where there is thus some evidence of slightly greater dissimilarity between the atoms, there is a distinctly greater hardening effect. Finally, in the case of the alloys of copper with manganese we have an alloy system still forming what appears to be an unbroken series of solid solutions, but the shape of the freezing point curve suggests that the system is one in which the series of solid solutions has only just been maintained and that a very slightly greater divergence between the component atoms would have led to the occurrence of a short range of duplex alloys. Here a still greater degree of hardening is found than in the systems previously mentioned.

---

* Beilby, "The Hard and Soft State in Metals," 'Phil. Mag.,' August, 1904; also 'Roy. Soc. Proc.,' A, vol. 76, p. 462 (1905), and 'Roy. Soc. Proc.,' A, vol. 79 (1907).

Still more striking evidence in support of the generalisation stated above is to be obtained from a consideration of alloy systems in which the range of solid solutions is definitely limited. Unfortunately, the ranges of solid solubility for alloys in true equilibrium have only been determined with sufficient accuracy in a comparatively few instances, while a detailed knowledge of the corresponding mechanical properties is also lacking in many cases. For the alloys of copper and for those of aluminium, however, adequate data are available. In regard to copper, the alloys with nickel and manganese have already been mentioned. In the well-known alloys of copper with zinc (brass, etc.), solid solutions can be obtained at the copper end of the series containing up to about 36 atomic per cent. of zinc (equivalent to about 37 per cent. of zinc by weight). The corresponding limit for aluminium in copper lies at about 14 atomic per cent. of aluminium (corresponding to about 7 per cent. by weight), while, for tin in copper, the limit lies near 6·7 atomic per cent. (equivalent to 12 per cent. by weight). In good correspondence with these figures are the facts concerning the hardening effects of these three metals on copper. It is difficult to give exact quantitative data, since tests on the alloys in a strictly comparable condition are not readily available, but there can be no doubt that the hardening effects of the three metals are distinctly in the inverse order of their limiting solid solubilities expressed in atomic percentages, as given above. Thus the well-known brass, containing 70 per cent. (by weight) of copper and 30 per cent. of zinc, is still very ductile. The ductility of the alloys of copper with aluminium decreases rapidly after an aluminium-content of 7 per cent. has been passed, while, in the case of the copper-tin alloys, an addition of 4 per cent. of tin (by weight) almost exhausts the ductility of the material. It appears probable, again to a first approximation, that the saturated solid solutions of copper containing any of the soluble metals will have approximately similar mechanical properties, independently of which metal has been used to bring about saturation. Since, however, the solubilities in some cases are high, the physical properties of the dissolved metal itself must affect such a conclusion, which is none the less broadly in conformity with the facts.

Similar evidence is to be obtained from a consideration of the alloys of aluminium with zinc and copper. In the aluminium-zinc series, solid solutions containing above 15 atomic per cent. of zinc (equivalent to a little under 30 per cent. of zinc by weight) can be obtained, while, in the aluminium-copper series, the limiting solid solubility lies in the neighbourhood of 1·5 atomic per cent. of copper (about 4 per cent. by weight). In conformity with this difference in solid solubility, we find that the hardening effect of copper on aluminium is very much greater than that of zinc. Actually, it is much more difficult, on account of slower diffusion, to obtain really saturated solid solutions of copper in aluminium than of zinc, but, when saturation with copper has been obtained, it is found that the physical properties of the resulting alloy (containing about 4 per cent. of copper by weight) are not very different from those of a solid solution of zinc in aluminium containing about 30 per cent. (by weight) of zinc.

Reference may further be made to the alloys of iron. The solid solubility of carbon in γ-iron is of the order of 2 per cent. by weight. In accordance with this figure, we find the enormous influence of carbon on the physical properties of iron, as compared with the effects of the highly soluble elements, such as nickel or silicon.

The Influence of Physical Conditions on the Velocity of Decomposition of certain Crystalline Solids.

By C. N. HINSHELWOOD, B.A., Fellow of Balliol College, and E. J. BOWEN, B.A., Balliol College.

(Communicated by Prof. F. A. Lindemann, F.R.S. Received February 25, 1921.)

The study of the velocity of chemical change has usually been confined to substances in the liquid or gaseous state. The temperature coefficient of the velocity is of special importance, since it can be used to calculate the "critical increment" or 'heat of activation" of the reacting molecule. In the case of reactions in solution the velocity is very largely influenced by the solvent, but the quantitative relationships do not conform to any simple law, owing to such factors as the formation of molecular complexes.

Reactions in the solid state have not been much studied, although in many respects the solid state is simpler, owing to the fixed arrangement of the atoms in a space lattice. The object of the experiments described in this paper was to ascertain whether simple quantitative relationships could be found for cases of "irreversible" decompositions of certain crystalline solids. The results show that in these cases the temperature coefficient is a complicated quantity, dependent on several physical factors; and that the heat of activation cannot be calculated from it. On the other hand, a simple relation is found connecting the rate of evolution of oxygen from solid solutions of potassium permanganate in potassium perchlorate and the heats of activation of the potassium permanganate in the various solid solutions.

It was shown in a previous paper* that potassium permanganate decomposes more slowly in solid solution in potassium perchlorate than in the pure state. According to the accepted view,† a molecule only decomposes when it possesses a sufficient excess of energy over the mean energy corresponding to the temperature of the system to which it belongs. This is the "heat of activation" and per gramme-molecule is identical with the quantity Q in Trautz's theorem, which represents the velocity constant k of a chemical change by an exponential function of the form

$$k = \chi e^{-Q/RT} \text{ where } \chi = \text{constant}.$$

If the heat of activation, Q, be modified by an amount q, then k will be changed to a value k' given by

$$ln\,\frac{k}{k'} = \frac{q}{RT}. \qquad (1)$$

---

* 'Phil. Mag.,' vol. 40, p. 569 (1920).

† *E.g.*, Lewis, 'J.C.S.,' vol. 109, p. 796 (1916).

As an element of the permanganate crystal structure—presumably the MnO_4 ion—becomes more stable when placed in a lattice consisting partly of perchlorate, it may be assumed that the heat of activation increases. If we assume that the difference in the heats of activation of potassium permanganate caused by a difference in its surroundings is equal to the heat given out when passing from one surrounding to another, the relative rates of decomposition in different solid solutions can be calculated from the heats of mixture of potassium perchlorate and permanganate in the solid state. These can be obtained from measurements of Sommerfeld, and calculation of the decomposition velocities from these data has been found to agree with direct measurement, as described in § 4.

Before this result could be obtained, however, a number of points connected with the dynamics of solid reactions had to be cleared up. In the first place, we have already shown (*loc. cit.*) that the decomposition by heat of crystalline substances, such as potassium permanganate and ammonium bichromate, takes place only on the surface of the crystals. Hence in any comparative experiments it is necessary to control the surface. It is found (§ 3) that the surface determining the velocity of reaction is considerably greater than the apparent surface deduced from the linear size of the crystals, since what is apparently a single crystal really consists of discontinuous aggregates.

The second complication relates to the fact that the surface may increase very greatly as the reaction proceeds, owing to the disintegration of large crystals. This was shown previously to occur with large crystals of potassium permanganate, the decomposition of which proceeds with a very marked acceleration. The acceleration largely disappeared when the crystals were ground up initially. The effect is illustrated by experiments on silver permanganate, described in § 1.

Another effect is sometimes observed in the early stages of the reaction in cases where the products form a coherent film on the surface of the crystal, the reaction being checked as soon as the surface has become completely covered with this film. We have therefore to consider not only the progress of the reaction with time, but its propagation into the interior of the crystal (§ 2). This complication, together with the fact that the effective surface may vary with temperature, renders the temperature coefficient useless for determining the heat of activation.

§ 1.—*Accelerated Decomposition (of Silver Permanganate).*

This substance evolves oxygen at a much lower temperature than the potassium salt, and measurements could be made at 110·0° C. The experimental method consisted in heating in a vapour bath about 0·5 grm. of the substance, contained in a small bulb and connected by a capillary tube with a gas-measuring burette. Small but well-formed crystals, whose mean longer axis was 0·25 mm. and whose shorter axis was on the average 0·027 mm., decomposed in a markedly accelerated manner, while when these were ground so that the mean diameter of the particles was about 0·005 mm. the initial rate was much increased and the acceleration very considerably diminished. The following results represent the mean of two concordant experiments in each case:—

Percentage of Total Oxygen evolved during Successive Periods of 30 minutes at 110·0° C.

	Crystals.	Powder.
0– 30 minutes	4·0	25·5
30– 60 "	4·0	40·0
60– 90 "	13·5	29·5
90–120 "	30·5	10·0
120–150 "	23·0	3·0

§ 2.—*Propagation of the Reaction from the Surface into the Interior.*

When a crystal decomposes, the change commences on the surface and is gradually propagated into the interior. It must be clearly understood that the process here referred to is an isothermal one, time having been allowed for the crystal to attain a uniform temperature throughout its mass. If the products of reaction do not form a coherent film, the change proceeds without hindrance, and, as already stated, may be strongly accelerated owing to the disintegration of the crystal. An experiment was previously described in which a large crystal of ammonium bichromate was allowed to decompose at 212° C. When the surface had become covered with a film of chromium oxide the gas evolution was checked. The effect is well seen by crushing a crystal in this state. With light pressure the fragments are all found to be covered with the black layer; further crushing shows the interior of the particles to be undecomposed.

Measurements made at a series of temperatures have now shown that as the temperature increases this retardation becomes less and less marked, and is succeeded by the usual type of accelerated reaction. The curves in fig. 1 show the results of a series of experiments with small crystals of ammonium bichromate. The substance was heated in glass vessels of about 10 c.c. capacity, connected with a mercury manometer, whereby the pressure developed was measured, the volume being maintained constant. Since water vapour is given off during the decomposition, the arrangement of the apparatus also ensured that the whole of the decomposition vessel was at the temperature of the vapour bath.

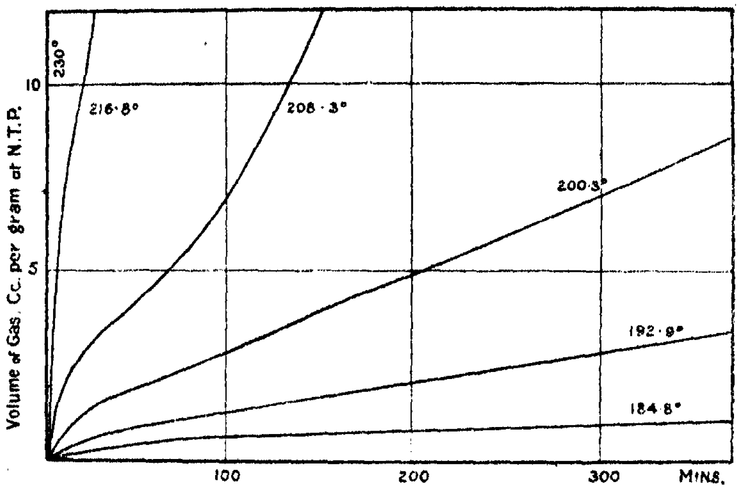


Fig. 1.—Decomposition of Ammonium Bichromate. Initial Stages.

If the times necessary for a given amount of gas evolution to occur at two different temperatures are compared, it is clear that the tendency of the retardation to disappear at the higher temperatures may give rise to what is apparently an abnormally high temperature coefficient. For example, the time required for the evolution of 1 c.c. per gramme at 200·3° C. was 18 minutes, while at 184·8° C. it was 330 minutes. Expressed in the conventional way, this means a temperature coefficient of 6·5 for 10°, which is much higher than is ever found for a homogeneous reaction. The abnormally high value results from the fact that not only does the reaction velocity increase as the temperature rises, but the reaction is also propagated more easily into the interior.

With potassium permanganate this effect is not observed, since the products are not firmly retained, but in the initial stages of the decomposition of the solid solutions in potassium perchlorate a retardation may be observed, because here the decomposition products are more or less firmly held by the surrounding structure. Fig. 2B illustrates the transition from the accelerated initial change of pure permanganate to the retarded decomposition of the dilute solid solutions. Solid solutions containing about 60 per cent. of potassium permanganate give an almost linear relationship between time and percentage decomposition. The results plotted in these curves were obtained with material ground as nearly as possible to the same state of subdivision.

Increase of temperature, as well as decrease in size of particles, renders the propagation into the interior easier, as may be seen from fig. 2A which represents the initial rates of decomposition of a single solid solution at various temperatures.

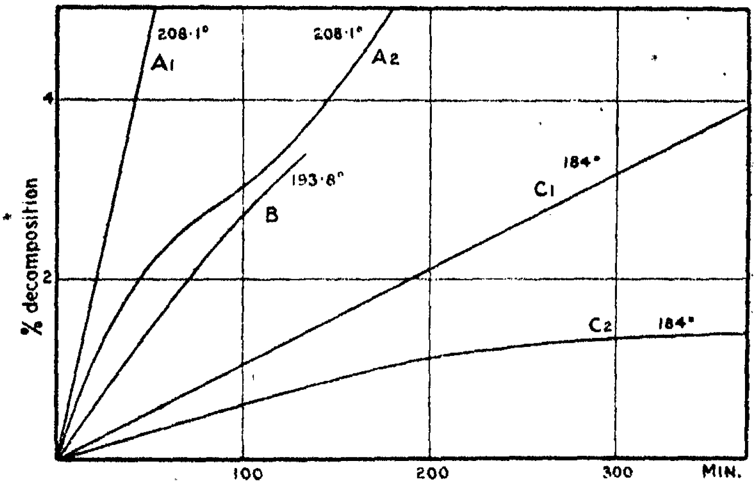


FIG. 2A.

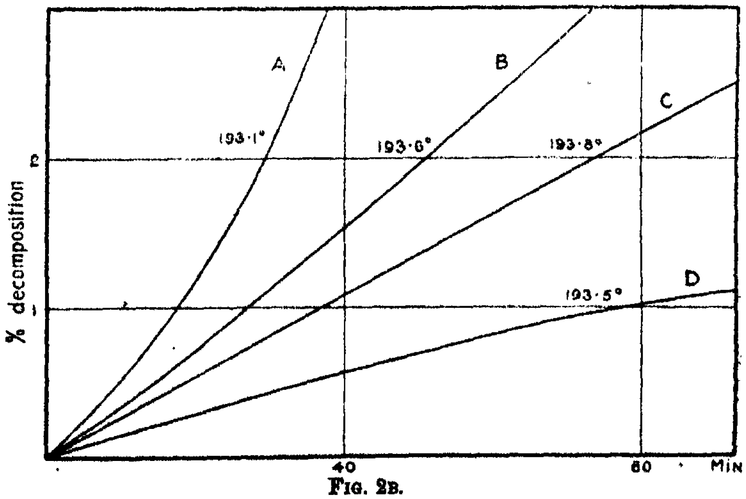


FIG. 2B.

FIG. 2A.—Solid Solution containing 60·5 per cent. $KMnO_4$. A1, B, C1, finely ground. A2, C2, small crystals.

FIG. 2B.—Solid Solutions of various Compositions. A, 100 per cent. $KMnO_4$. B, 76·5 per cent. $KMnO_4$. C, 60·5 per cent. $KMnO_4$. D, 24·7 per cent. $KMnO_4$.

§ 3. *Influence of the Size of Particles on the initial Rate of Reaction.*

While it is quite clear qualitatively that the rate of reaction increases with the degree of subdivision, the exact relation between rate and crystal size requires further consideration. If a given mass of substance the mean linear magnitude of whose particles is a be subdivided into particles whose mean linear magnitude is b, then the surface is increased in the ratio a/b. But experiment shows that the velocity of reaction does not increase in this ratio, but in a considerably smaller one. This implies that the surface which is effective in determining the rate of reaction is considerably greater than the apparent surface as deduced from the linear size of the particles. What appears to be an individual crystal is really composed of aggregates, some of which are sufficiently loosely held together to allow decomposition to take place between them. This is in accord with the results derived from the consideration of tensile strength, which is always much lower than would be expected if the material were continuous. From this it is concluded that the tensile strength represents the force between small aggregates.

The following is a typical result:—

Solid solution of $KMnO_4$ in $KClO_4$,

60·6 molecular percentage.

Mean diameter of particles.	Initial rate of decomposition at 184·3 C.
0·38 mm.	0·0054 per cent. per min.
0·026 mm.	0·0106 per cent. per min.

The ratio of the velocities is 1·96.

At 208·1° C. the ratio was found to be 2·13. Taking the mean as 2·05, it is seen that this is considerably less than the inverse ratio of the linear sizes $0{\cdot}38/0{\cdot}026 = 15$.

§ 4.—*The Evolution of Oxygen from Isomorphous Mixtures of Potassium Permanganate and Perchlorate. Dependence of Reaction Velocity on the Concentration of Potassium Permanganate.*

It may be stated at the outset that at the temperature of the experiment only the potassium permanganate undergoes measurable decomposition. Nevertheless, the stability of the permanganate elements of the crystal structure is influenced by the presence of the perchlorate elements. It is, therefore, of interest to determine the velocity of oxygen evolution from solid solutions of various compositions.

It is clear that to find out how the stability of the potassium permanganate crystal element depends on its environment, the decomposition should only be measured in its initial stages, since if it be allowed to proceed to a more advanced stage complications may occur, due to the varying facility with which the change is propagated into the interior of the crystals of different compositions. In the different experiments of a series it is necessary to use crystals which possess the same surface. This is difficult, since the apparent surface is not a real index of the true surface. One must, however, employ crystals of the same apparent size and regard the results as partly statistical.

The determination of the initial velocities was attempted by two methods. The first was to follow the decomposition to about 5 per cent. and determine the initial rate directly. The second was to follow the decomposition over its whole course and to extrapolate. But the changes in curvature near the commencement are so marked that the extrapolation must be regarded as rather arbitrary.

The experimental method consisted in heating in a vapour-bath about 0·5 grm. of the substance contained in a small glass bulb, and connected by a capillary tube to a mercury manometer in the first series and to a small gas-burette in the second. The oxygen evolved was thus measured, either by increase in pressure in the apparatus, or by direct reading of its volume under atmospheric pressure.

The following results were obtained. (The figures represent initial velocities in per cent. per minute):—

Temp.	Mol. p.c. $KMnO_4$. 100 p.c.	Temp.	Mol. p.c. $KMnO_4$. 76·5 p.c.	Temp.	Mol. p.c. $KMnO_4$. 60·5 p.c.	Temp.	Mol. p.c. $KMnO_4$. 58·5 p.c.
148·0	0·00185	182·3	0·0185	184·3	0·0106	182·3	0·0038
183·4	0·0295	193·2	0·042	193·8	0·0270	192·7	0·0138
192·3	0·0550			208·1	0·103	206·2	0·056
193·1	0·053						
Value deduced for 200° 0·091.		Value deduced for 200° 0·070.		Value deduced for 200° 0·048.		Value deduced for 200° 0·029.	

Temp.	Mol. p.c. $KMnO_4$. 46·2 p.c.	Temp.	Mol. p.c. $KMnO_4$. 33·5 p.c.	Temp.	Mol. p.c. $KMnO_4$. 24·7 p.c.	Temp.	Mol. p.c. $KMnO_4$. 24·3 p.c.
193·5	0·085	193·2	0·024	207·2	0·056	229·3	0·39
				230·2	0·38		
Value deduced for 200° 0·059.		Value deduced for 200° 0·042.		Value deduced for 200° 0·031.		Value deduced for 200° 0·034.	

The effect of temperature varies from one solid solution to another. This must be attributed to a greater or lesser loosening of the crystal aggregates as the temperature rises. For this and the other physical reasons previously discussed, it is of little value to calculate the "heat of activation" from the temperature coefficient.

From these results the following series of relative values at 200° C. is found by plotting the logarithm of the initial rate against temperature.

Molecular percentage $KMnO_4$.	Relative initial velocity.
100·0	100
76·5	77
60·5	53
58·5	32
46·2	65
33·5	46
24·7	34
24·3	37

In fig. 3, the curve B is drawn in the most probable position through these points. Considerable deviations from the mean are to be expected owing to variations in the effective surface.

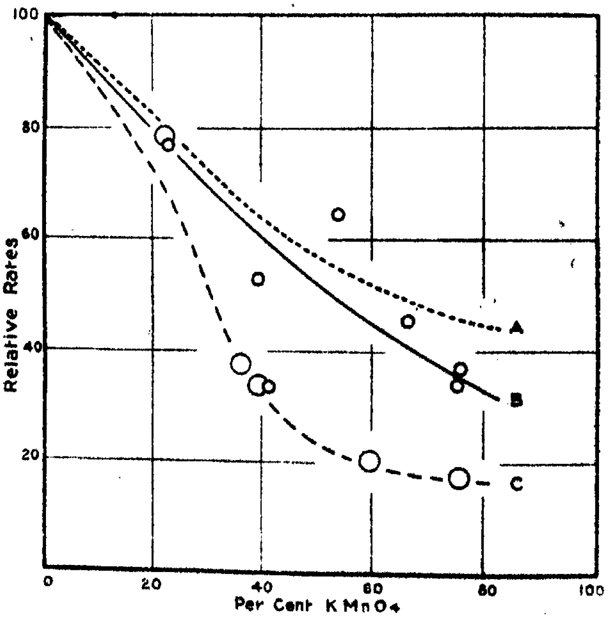


Fig. 3.—Relative Velocities of Decomposition of Solid Solutions of $KMnO_4$ in $KClO_4$ at 200°. A, Initial Rates calculated from heat of mixture. B, Initial Rates observed. C, Rates estimated arbitrarily from times required for 20 per cent. decomposition.

In the second series the whole course of the change was followed at 240° C. small crystals selected by sieving being used. The relative velocities as deduced from the times required for 20 per cent. decomposition are given below and in the curve C in fig. 3. This curve is steeper than A owing to the physical cause already discussed.

Molecular percentage $KMnO_4$.	Relative velocity.
100·0	100
76·5	79
63·0	38
60·5	34
40·8	21
24·7	18

From the curve of initial velocities (curve B) in fig. 3 we can test equation (1) according to which if the heat of activation is increased by q the velocity constant is reduced from k to k' where

$$ln\ \frac{k}{k'} = \frac{q}{RT}.$$

Putting initial velocities proportional to k we have, for an *equimolecular mixture* at 200°,

$$\frac{k}{k'} = \frac{100}{52},$$

whence
$$ln\ \frac{100}{52} = \frac{q}{1·98 \times 473},$$

$\therefore$ $q = 618$ calories per gramme molecule $KMnO_4$,

or 309 calories per gramme molecule of the equimolecular mixture.

This has been assumed equal to the heat of solid mixture.

This heat of solid mixture can be determined directly from the heats of solution of potassium permanganate, perchlorate and the mixed crystals. E. Sommerfeld* has determined these accurately, and from his results it is clear that there is a small evolution of heat when mixing takes place in the solid state. For 1 gramme molecule of the equimolecular mixture it amounts to 260 calories. Since this represents the difference between heats of solution which amount separately to about ten thousand calories, better agreement cannot be expected, with the value estimated above. Moreover, these measurements were made about 180° below those of the reaction velocities, so that a comparison implies that the specific heats of the solid solution can be calculated additively from those of the components.

---

* 'Neues Jahrbuch für Mineralogie,' Beilage Band XIII, p. 443 (1900).

From a curve of Sommerfeld's results the following heats of mixture per gramme molecule of potassium permanganate are deduced:—

	Per cent.	Per cent.	Per cent.	Per cent.
$KMnO_4$	80	60	50	40
$KClO_4$	20	40	50	60
Calories	159	403	520	600

Using equation (1), we may calculate from these figures the relative rates of decomposition of potassium permanganate in solid solutions of various compositions. These are shown in curve A in fig. 3, and it will be seen that the theoretical curve lies very close to the curve B, representing the experimental results. From this it is clear that the lowering in velocity is explained by the change in heat of activation being equal to the heat of solid mixture.

The results described in this paper once more seem to confirm the view that no distinction can be drawn between physical and chemical forces. The chemical forces between the atoms of solid substances are profoundly modified by physical conditions, such as proximity to the surface, and by the presence of neighbouring atoms in the space lattice. Moreover, the decrease in potential energy attending the physical process of solid mixture leads directly to increased chemical stability in a way which can be calculated approximately from the heat of mixture. It will be interesting to see whether similar agreement can be found in the case of other isomorphous mixtures, including those with a negative heat of formation.

We have pleasure in thanking Brig.-Gen. H. Hartley for his interest in this work.

Heats of Combustion and Formation of Nitro-Compounds.
Part I.—*Benzene, Toluene, Phenol and Methylaniline Series.*

By W. E. Garner and C. L. Abernethy.

(Communicated by Sir Robert Robertson, F.R.S. Received December 15, 1920.)

Introduction.

The study of the explosive properties of organic nitrobodies is not complete without a knowledge of their heats of formation, for these constants are closely related to their stability and sensitiveness to impact and heat. The heats of formation also form the basis of any calculations of the energy liberated in the detonation of high explosives. The heats of detonation of balanced explosives such as a mixture of ammonium nitrate and trinitrotoluene arranged for complete combustion, can readily be determined from the heats of formation. When there is insufficient oxygen in the explosive for the complete combustion of the hydrogen and carbon it is more difficult to calculate the heat of detonation, especially as the composition of the products is dependent on the external work done by the gases, but from a knowledge of the constants of gaseous and heterogeneous equilibria it is often possible to obtain values with some degree of accuracy. The preliminary work in the production of new explosives can often be considerably shortened by such calculations.

The heats of combustion of some of the members of the toluene, benzene and phenol series have been determined by various workers, but there are many gaps in the series, and the agreement in certain cases is unsatisfactory. These results are summarised in Table I.

In order to amplify these data, it was decided to carry out an accurate estimation of the heats of combustion and formation of the nitro-compounds in the toluene, benzene, phenol and methylaniline series, and to calculate both the heats due to the entrance of a nitro-group into members of the series and the heats of nitration.

The toluene series was investigated in greater detail, since all of the isomers in this series were available. The constants of only the principal members of the other series have been determined.

Apparatus and Methods of Manipulation.

The calorimetric system consisted of a Berthelot–Mahler–Kroeker bomb placed in a nickel-plated copper vessel containing 1200 c.c. of water; this vessel was surrounded by another larger similar vessel, the whole standing

Table I.—Heats of Combustion of Nitro-compounds as given by Previous Observers.

Series.	Substance.	Mol. wt.	Heat of combustion.			Heat of formation.	Observer.
			Grm. cal. per grm.	Kgrm. cal. per mol.			
				Const. vol.	Const. press.		
Toluene	*m*-Nitrotoluene	137	—	889·4	889·6	11·0	Swarts.
	Trinitrotoluene—						
	2 : 4 : 6	227	—	—	—	7·7	From Bichel's figures.
	2 : 4 : 6	227	—	—	—	5·9	Koehler.
	Crude T.N.T.	227	—	—	880·8	1·8	Will.
Benzene	Nitrobenzene	123	—	734·8	734·7	3·9	Swarts.
	m-Dinitrobenzene	168·05	—	695·3	695·1	8·6	Swarts.
	o-Dinitrobenzene	168·05	4194	704·8	703·7	—	B. Matignon.
	m-Dinitrobenzene	—	4155	698·3	697·3	—	B. Matignon.
	p-Dinitrobenzene	—	4146	696·7	695·6	—	B. Matignon.
	1 : 3 : 5-Trinitrobenzene	213·05	3126	666·1	664·0	—	B. Matignon.
	1 : 2 [illegible] Trinitrobenzene	—	3195	680·8	678·6	—	B. Matignon.
Phenol	*o*-Nitrophenol	139·05	4954	688·8	688·4	—	Matignon, Deligny.
	p-Nitrophenol	—	4960	689·7	689·3	—	Matignon, Deligny.
	o-Nitrophenol	—	—	688·5	688·0	50·1	Swarts.
	m-Nitrophenol	—	—	684·8	688·4	53·7	Swarts.
	p-Nitrophenol	—	—	688·6	688·2	50·0	Swarts.
	Picric acid	229·05	—	—	618 4	50·9	Sarrau and Vielle.
	Picric acid	—	—	—	—	50·2	Koehler.
Aniline	*o*-Nitraniline	138	—	766·2	765·9	6·9	Swarts.
	m-Nitraniline	—	—	766·6	766·3	6·5	Swarts.
	p-Nitraniline	—	—	760·5	760·2	12·3	Swarts.
	Trinitraniline	228	—	—	—	15·2	Koehler.
	Tetryl	287	—	—	—	40·8	Koehler.

within a copper water jacket. A thick felt covering protected the whole system from outside influences, and closely fitting cardboard lids prevented air currents from entering. The air space in the inner calorimeter was reduced to a minimum to avoid convection currents, and the volume of water was as small as was consistent with good stirring.

The water equivalent was determined by means of the secondary standards, naphthalene and benzoic acid. There is, however, some confusion in the literature as to the best values for the mechanical equivalent of heat and the secondary standards. The three main secondary standards, naphthalene, benzoic acid and cane sugar, have in turn been criticised by different authors. Their usefulness appears to be in the order given. The values of the heats of combustion of these substances are still uncertain to the extent of two to three parts per thousand. This is due chiefly to the different values taken for the mechanical equivalent of heat by German and by American workers, but even the ratios of the heats of combustion of different substances obtained by different workers are unsatisfactory. The determinations of Roth, Fischer and Wrede, Wrede and the American Bureau of Standards are based on electrical methods which have been carried out with great care. These together with data from other authors are summarised below :—

Table II.—Heats of Combustion of Secondary Standards by various Observers.

No.	Substance.	Cal./grm.	Joules per grm.	Observer.
2	Naphthalene	9624 (Roth.)	40314	Wrede, 'Zt. Phys. Chem.,' vol. 75, p. 92 (1910).
1	”	—	40384	Fischer, 'Sits. Akad. Wiss. Berlin,' vol. 15 (1904).
—	—	—	40350	Wrede, Swientoslawski, 'J. A. C. S.,' vol. [illegible], p. 2598 (1917).
5	”	9612	—	Di[illegible], 'Bureau of Standards Sci. [illegible]0.
6	”	9613	—	[illegible]ki, 'J. A. C. S.,' p. 2598 (1915).
3	”	9636	[illegible]	[illegible]n.,' vol. 407, p. 127
4	”	9628	[illegible]	[illegible] Prakt. Chem.,' (II), [illegible]9).
2	Benzoic acid	631[illegible]	[illegible]	[illegible]
1	” ”	[illegible]	[illegible]	[illegible]hys. Chem.,' vol. 69, [illegible]
5	” ”	[illegible]	[illegible]	[illegible]
6	” ”	[illegible]	[illegible]	[illegible]lawski.
4	”	[illegible]	[illegible]	[illegible]mann.

The [illegible] of combustion is not altogether sat[illegible] agreement between absolute values or between the ratios of the heats of combustion of the two substances as has been pointed out by Swientoslawski (*loc. cit.*).

The ratios for naphthalene and benzoic acid have been recalculated and are given in the following Table :—

Table III.—Ratio of Heats of Combustion of Naphthalene and of Benzoic Acid.

No.	Author.	$\dfrac{\text{Heat of combustion of naphthalene}}{\text{Heat of combustion of benzoic acid}}$.	Date.
1	Fischer and Wrede	1·5254	1904
2	Wrede	1·5238	1910
3	Roth	1·5235	1915
4	Stohmann	1·5229	1889
5	Dickinson	1·5202	1915
6	Swientoslawski	1·5232	1917
7	Research Department, Woolwich	1·5240	1918

Of these values the highest (1) must be regarded as superseded by (2), and (5) ought to be eliminated on the showing of Swientoslawski. The mean of the four remaining values (1·5232) is the same as that determined by Swientoslawski, and it seems advisable to accept this value provisionally. The German author Roth (see Landolt and Börnstein's Tables) has assumed the mechanical equivalent of heat to be 4·189 ergs, the value of the Reichsanstalt, Berlin, while (5) and (6) in Table II are based on the figure 4·182 ergs, the value of Harvey and Davis, Harvard.

Since the determination by the American Bureau of Standards agrees with that of Swientoslawski and is more recent than that of Roth and Wrede, the 15° value of 9613 cal./grm. may be provisionally accepted for the heat of combustion of naphthalene, and from the above consideration of the ratios the most reliable value for benzoic acid is 6311 for the 15° calorie. To make the results in this paper directly comparable with those of Richards* the 18° calorie is taken in this work as the unit, making the heat of combustion of naphthalene 9622 cal./grm. and of benzoic acid 6317 cal./grm. These values are employed for the determination of the water equivalent of the system.

The mean value obtained by calibration with benzoic acid was 1544·8 cal./grm. The ratio of the heats of combustion of naphthalene

---

[*] 'J. A. C. S.,' vol. 39, p. 341 (1917). Since this paper was written Richards and Davis ('J. A. C. S.,' vol. 42, p. 1599 (1920)) have re-determined the heats of combustion as follows :—Benzoic acid, 6320 ; naphthalene, 9614 (18° calories per grm.). These results show that there is still uncertainty in the values of the secondary standards.

Table IV.—Water Equivalent of System, by Combustion of Naphthalene.

Substance.	No. of Beckmann thermometer.	Weight.	Temp. rise.	Water equivalent.	
		grm.	° C.		
Sample 2— M.P. 80·3–80·5° C.	No. 2	0·4822	3·004	1544·4	1543·9 grm.
	"	0·4989	3·108	1544·4	
	1329 No. 14	0·3302	2·060	1542·4	
	"	0·3201	1·994	1544·5	
	"	0·3985*	2·499	1542·0	
	P.T.R. 59334	0·3952*	2·4745	1544·3	
	"	0·4055	2·525	1545·3	

* In the two cases marked with an asterisk 0·0046 grm. cotton-wool (heat of combustion 4020 calories per gramme) was used to facilitate ignition of the pellet of naphthalene.

and benzoic acid was found to be 1·5240, which differs by 8 in 15,000 from the mean ratio in Table III.

Owing to the variations in the published heats of combustion of benzoic acid, and the difficulties which occur in its purification, the value obtained with naphthalene, 1544, is considered the more accurate, and so was adopted as the water equivalent of the system.

All the solid substances were burnt in the form of pellets, and were fired electrically by platinum wire of 0·004 inch diameter, and a battery of 12 volts. In order to avoid overheating of the top of the bomb and too rapid combustion of the substance, the weight taken was such as to give a rise in temperature of not more than 1·5° C. and 2° C. The bomb was twice filled with oxygen to a pressure of 5 atmospheres and emptied, before finally filling to a pressure of 20 atmospheres of oxygen. The percentage of atmospheric nitrogen was thus greatly reduced, so that the nitric acid which did form was mainly due to the nitrogen in the explosive. The rate of stirring was kept constant at about seventy strokes per minute.

In the case of liquids such as nitrobenzene, a modification of Richards' method* was used. At first, the nitrobenzene was placed in a thin glass bulb at the bottom of the crucible, and ignited with a tuft of cotton-wool and paraffin-wax, as in Richards' method. Residues of carbon were left on the top and the walls of the bomb, and, although there was no smell of nitrobenzene, the heats of combustion were irregular. Fairly satisfactory results were obtained if the unconfined liquid lay at the bottom [illegible] crucible and Richards' method of ignition employed. Wi[illegible] firing, the nitrobenzene was completely observed in the products of combustion. All the liquids investigated were fired by this method.

The corrections applied to obtain the heats of combustion at constant pressure were:—

(i) Radiation correction.
(ii) Correction for calibration of the thermometer.
(iii) Correction for the heat of formation of the nitric acid produced.
(iv) Correction of the gases to constant pressure.

(i) The maximum temperature being nearly reached at the end of the first minute, a short method was found to give the radiation correction sufficiently accurately.

(ii) The thermometer used was at first Beckmann No. 2, N.P.L. 13, and, later, Beckmann 1329, N.P.L. 14, both of which had been standardised to 0·002° C. The water equivalent was not appreciably altered by changing the thermometer.

(iii) The nitric acid produced during the combustion was estimated, and correction applied for the heat of formation of nitric acid.

$$5O_2+2N_2+2H_2O+Aq = 4HNO_3Aq+4\times 14{\cdot}9 \text{ kgrm.-cal.}$$

The correction (due to nitric acid) is about 1 to 1·5 kgrm.-cal./mol. for mononitro-compounds, and increases to 1·7 to 2·6 for trinitro-compounds. The amount of nitric acid formed was somewhat variable, probably owing to irregularities in the rate of combustion of the nitro-compound. It was therefore necessary to determine and allow for the nitric acid produced in every experiment.

(iv) To convert the molecular heat of combustion at constant volume to the molecular heat of combustion at constant pressure, 0·54 + 0·0021 kgrm.-cal. is subtracted for every mol. of gas formed, and added for every mol. of oxygen used up in the combustion of one mol. of substance. It is assumed that the whole of the water formed becomes liquid, and the error thus introduced does not affect the fifth figure in the value of the heat of combustion.

The value 0·58 kgrm.-cal., corresponding to a temperature of 20° C., has been used.

Purification of Nitro-compounds.

The substances in the toluene and benzene series were supplied by the organic laboratory of the Research Department, and the solids had been purified until a constant melting point was obtained. The melting points and some boiling points are given in Table V.

Table V.

Melting Points of Nitro-compounds.

Substance.	Melting point.	Substance.	Melting point.
	° C.		° C.
Paranitrotoluene	52·1	*m*-Dinitrobenzene	90·0–90·1
2 : 3-Dinitrotoluene	59·6	Trinitrobenzene (Sym.)	122·1–122·4
2 : 4-Dinitrotoluene	70·5	*o*-Nitrophenol	45·1
2 : 5-Dinitrotoluene	51·2	*p*-Nitrophenol	113·1–113·9
2 : 6-Dinitrotoluene	64·3	2 : 4-Dinitrophenol	111·2–111·8
3 : 4-Dinitrotoluene	58·9	Picric acid	120·6–121·1
3 : 5-Dinitrotoluene	92·8	Paranitromethylaniline	151·6–151·9
a-Trinitrotoluene	81·1	Dinitromethylaniline	177·6–177·9
β-Trinitrotoluene	110·9–111·0	Trinitromethylaniline	110·4–110·8
γ-Trinitrotoluene	102·7–103·3	Tetryl	128·7–129·1
δ-Trinitrotoluene	132·0–132·3	Methyl tetryl	99·9–101·1
ε-Trinitrotoluene	95·8		
ζ-Trinitrotoluene	110·0		

Boiling Points of Nitro-compounds.

Substance.	Boiling point.
o-Nitrotoluene	222·3–222·7 (772 mm.)
m-Nitrotoluene	234·0–235·4 (772 mm.)
Nitrobenzene	211·1–211·3 (771 mm.)
Monomethylaniline	195·1–195·6

Heat of Combustion.

The collected results of this investigation are given in Tables VI to XIII. The values for benzene and toluene are Richards'* and are based on the 18° calorie, and calculated on the weight of substances *in vacuo*. The value for phenol is Berthelot's.† All other values were obtained by experiment in the same calorimetric system and calculated on the weight of substance in air. The heats of combustion of isomeric forms are found to differ considerably, the differences being altogether outside the limits of error for any one isomer.

An analysis of the results for each substance shows some small divergence for any given substance, and the variation was not diminished by an improvement in the calorimetric system. The explanation of this probably lies in the great rapidity with which nitro-compounds burn in oxygen. For example, it was found dangerous to burn more than 0·5 grm. of tetryl in the bomb, as spitting caused injury to the porcelain lining. A greater divergence may be expected among the more highly nitrated bodies, and this is borne out by an analysis of the Table.

---

* 'J. A. C. S.,' vol. 37, p. 1919 (1915).

† 'Ann. Chim. Phys.,' 6, vol. 10, p. 433 (1887).

No member of the methylaniline series, except tetryl, has given trouble in this respect, and the widest variation amongst the results for this series amounts to about 6 in 4500.

Heat of Formation.

The values for the heats of formation are calculated from the heats of combustion at constant pressure, both on the diamond and on the amorphous carbon basis, using the heats of formation given below.

Substance.	Heat of formation.	Observer.
Water	68·3 kgrm.-cal./mol.	Thomsen.
Carbon dioxide	97·0 kgrm.-cal./mol. (C. amorphous)	Thomsen.
	94·3 kgrm.-cal./mol. (C. diamond)	Berthelot.

The heats of formation of the nitro-compounds from amorphous carbon have been employed in calculations of the heats of detonation of high explosives, more especially as the deposited carbon from many high explosives deficient in oxygen is in the amorphous form. The heats of formation given in Tables VII, IX, XI, and XIII refer to that phase of the substances which exists at 15° C. In fig. 1 the heats of formation (amorphous carbon) are plotted against the number of nitro-groups in the molecule. The heats of formation in these cases are corrected to the solid state from values of the heats of crystallisation given in the Tables.

Heat of Entry of Nitro-group.

The heat of entry of the nitro-group into the molecule is given by the difference between the heats of formation of the substance and that of the next higher stage of nitration, and values obtained in this way for toluene are to be found in Table XIV.

Heat of Nitration.

The heats of nitration (Tables VII, IX, XI, and XIII) are calculated from the heats of formation in accordance with the equation:—

$$\begin{array}{ccc} C_7H_8 + HNO_3(Aq) & = & C_7H_7NO_2 + H_2O + 25{\cdot}3 \text{ kgrm.-cal.} \\ +15{\cdot}1 + 48{\cdot}8 \text{ kgrm.-cal.} & & +20{\cdot}9 + 68{\cdot}3 \text{ kgrm.-cal.} \end{array}$$

The values thus represent the heat which would be evolved in the nitration of the substances with aqueous nitric acid.

Fig. 2 shows the effect of the number of nitro-groups in the molecule on the heat of nitration for each stage, all substances being in the solid form. In the case of the toluenes, fig. 3 gives the heats of nitration in another form, which illustrates the effects of nitration in the ortho-, meta-, and para-positions to existing nitro-groups.

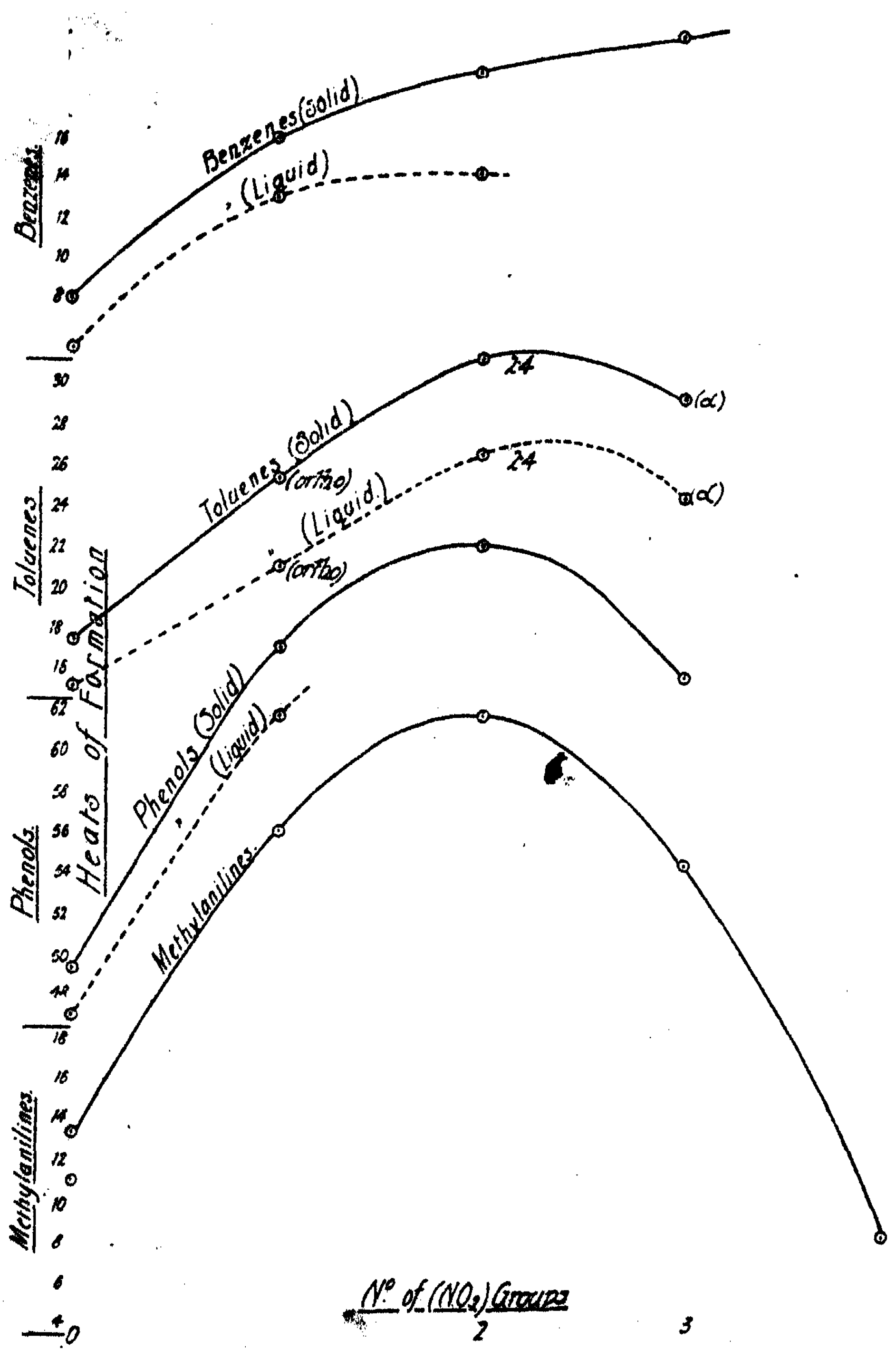

Benzenes (Solid)
" (Liquid)
Toluenes (Solid)
(ortho)
2·4
(α)
" (Liquid)
(ortho)
2·4
(α)
Phenols (Solid)
" (Liquid)
Methylanilines.
Heats of Formation
Benzenes.
16
14
12
10
8
Toluenes
30
28
26
24
22
20
18
16
Phenols.
62
60
58
56
54
52
50
48
Methylanilines.
18
16
14
12
10
8
6
4
No. of (NO2) Groups
0
2
3

FIG. 1.

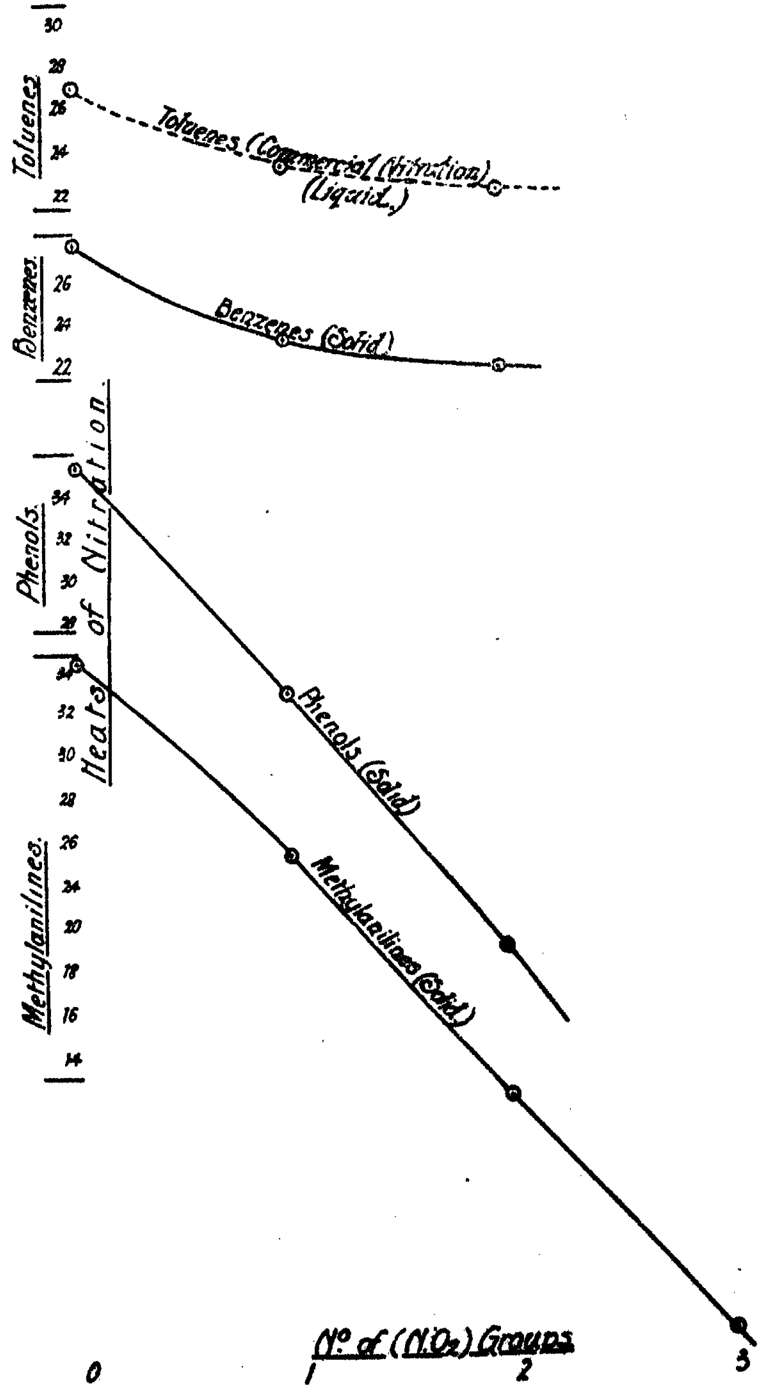
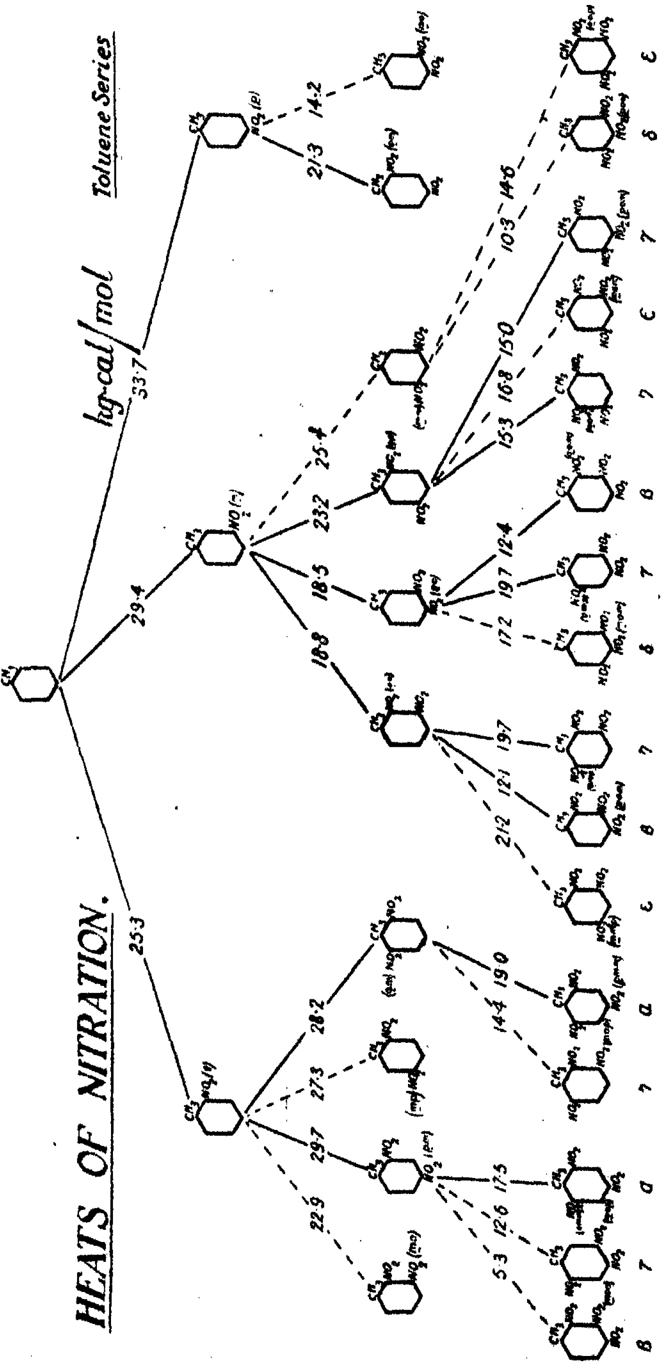

Toluenes
30
28
26
24
22
Toluenes (Commercial (Nitration)
(Liquid.)
Benzenes.
26
24
22
Benzenes (Solid)
Heats of Nitration.
Phenols.
34
32
30
28
Phenols (Solid)
Methylanilines.
34
32
30
28
26
24
20
18
16
14
Methylanilines (Solid)
No of (NO2) Groups
0
1
2
3

FIG. 2.



Unbroken lines indicate that direct nitration can be brought about, broken lines that this is not the case, the new compounds being obtained by indirect means.

Letters in brackets indicate position of entering group relative
(a) to CH₃ group if underlined
(b) to NO₂ group if underlined.
N.B. Corrections for heat of crystallisation not included.

FIG. 3.

Heat of Crystallisation.

Some of the heats of crystallisation of members of the benzene, toluene and phenol series are known, and these are included in the following tables; but no data are available as to heats of crystallisation of isomeric nitro-bodies.

In order to determine to what extent the differences in the heats of formation of the isomeric trinitrotoluenes are affected by differences in the heats of crystallisation, the heats for the trinitrotoluenes were compared with Tammann's* value (4·88) for α-trinitrotoluene by a method involving measurement of rate of cooling.

The heats of crystallisation of β- and γ- T.N.T. were found to be 5·0 and 5·4 cal./mol. respectively; it is therefore considered that the large differences observed in the heats of combustion of the isomeric nitrotoluenes cannot be accounted for by differences in the heats of crystallisation.

Results.

(a) *Toluene Series.*—The nitrotoluenes form a complete series of isomeric bodies and are very suitable for an investigation into the relationship between the internal energy of isomeric bodies and their chemical constitution. The heats of combustion of the isomers of mono-, di- and tri-nitrotoluenes given in Table VI indicate that the internal energies of the various isomers vary considerably; thus the total variation in the dinitrotoluenes is 7·1 kgrm. cal./mol. and for trinitrotoluenes 12·2 kgrm.-cal./mol. The heats of combustion of the trinitrotoluenes run generally in line with their figures of insensitiveness and stability (Table VII) so that the isomer with the greatest internal energy (strain) is the most sensitive. Thus, β-trinitrotoluene is more sensitive than α-trinitroluene, and its heat of combustion is found to be higher.

Table VI.—Heats of Combustion of Toluene Series.

Substance.	Mol. wt.	Cal./grm. in oxygen uncorrected for nitric acid. (Average.)	Kgrm.-cal./mol. in oxygen uncorrected for nitric acid. (Average.)	Kgrm.-cal./mol. at constant volume corrected for nitric acid.
Toluene	92	—	—	935·9 (Richard's value).
Mononitrotoluene	137			
Ortho (liquid)	—	6569 6550 6548 *6567 } 6559	900·0 897·4 897·1 *899·7 } 898·2	897·0
Meta (liquid)	—	6528 6520 6533 } 6527	894·4 893·2 895·1 } 894·2	892·9

---

* Tammann, 'Zeit. Phys. Chem.,' vol. 85, p. 283 (1913).

Table VI.—*continued.*

Substance.	Mol. wt.	Cal./grm. in oxygen uncorrected for nitric acid.	(Average.)	Kgrm.-cal./mol. in oxygen uncorrected for nitric acid.	(Average.)	Kgrm.-cal./mol. at constant volume corrected for nitric acid.
Mononitrotoluene (*cont.*)						
Para	—	6490	6494	889·1	889·7	888·6
		6490		889·1		
		6496		889·9		
		6501		890·6		
Dinitrotoluene	182					
2 : 3	—	4743	4736	863·2	861·9	860·5
		4734		861·6		
		4730		860·0		
2 : 4 (two samples)	—	4711	4700	857·4	855·4	853·7
		4697		854·8		
		4700		855·4		
		4701		855·5		
		4693		854·0		
2 : 5	—	4711	4713	857·3	857·7	856·1
		4710		857·2		
		4718		858·6		
2 : 6 (two samples)	—	4717	4710	858·4	857·3	855·2
		4715		858·2		
		4713		857·7		
		4705		856·4		
		4701		855·6		
3 : 4 (two samples)	—	4736	4739	861·9	862·4	860·8
		4736		861·9		
		4741		862·9		
		4741		*862·8		
3 : 5	—	4699	4699	855·2	855·2	853·9
		4695		854·5		
		4703		855·9		
Trinitrotoluene	227					
α (2 : 4 : 6)	—	3631	3630	824·2	823·9	822·5
		3622		822·2		
		3629		823·7		
		3638		825·8		
		3628		823·6		
β (2 : 3 : 4)	—	3687	3684	837·2	836·3	834·7
		3679		835·2		
		3689		836·4		
γ (3 : 4 : 6)	—	3652	3654	828·9	829·4	827·4
		3654		829·6		
		3654		829·6		
δ (3 : 4 : 5)	—	3667	3667	832·4	832·3	829·9
		3667		832·5		
		3666		832·1		
ε (2 : 3 : 5)	—	3643	3648	826·9	828·0	825·6
		3645		827·4		
		3655		829·6		
ζ (2 : 3 : 6)	—	3647	3647	828·0	828·0	827·1
		3647		828·0		
		3643		827·0		
		3645		827·4		
		3653		829·2		

* Air not driven out of bomb. Correction made independently.

Table VII.—Summary of Heats of Combustion, Formation and Nitration per Molecule, and Figures of Insensitiveness.

Substance.	Mol. wt.	Heat of combustion (kgrm.-cal./mol.).		Heat of formation (kgrm.-cal./mol.).		Heat of nitration (kgrm.-cal./mol.).			Heat of crystallisation (kgrm.-cal./mol.).	Gas evolution at 120° C. in c.c. per 40 hours.	Figure of insensitiveness (picric acid = 100.).
						One NO_2 group inserted in the following positions relative to existing NO_2 groups.					
		Const. vol.	Const. press.	C. amorph.	C. diamond.	o-position.	m-position.	p-position.			
Toluene	92	935·9	937·1	+15·1	− 3·8	+25·3	+29·4	+33·7	2·4		
Mononitrotoluene—											
Ortho (liquid)	137	897·0	897·2	+20·9	+ 2·0	2 : 3+22·9	2 : 4+29·7 2 : 6+28·2	2 : 5+27·3			
Meta (liquid)	—	892·9	893·1	+25·0	+ 6·1	2 : 3+18·8 3 : 4+18·5	3 : 5+25·4	2 : 5+23·2	4·3		
Para	—	888·6	888·8	+29·3	+10·4	3 : 4+14·2	2 : 4+21·3	—			
Dinitrotoluene—											
2 : 3	182	860·5	859·6	+24·3	+ 5·4	2 : 3 : 4+12·1	2 : 3 : 5+21·2	2 : 3 : 6+19·7 2 : 3 : 5+21·2			
2 : 4	—	853·7	852·8	+31·1	+12·2	3 : 4 : 6 12·6 2 : 3 : 4 5·3	2 : 4 : 6 17·5	3 : 4 : 6 12·6	4·80		
2 : 5	—	856·1	855·2	+28·7	+ 9·8	2 : 3 : 5 16·8 3 : 4 : 6 15·0 2 : 3 : 6 15·3	2,3 : 5 16·8 3 : 4 : 6 15·0 2 : 3 : 6 15·3				
2 : 6	—	555·2	854·3	+29·6	+10·7	2 : 3 : 6 14·4	2 : 4 : 6 19·0	2 : 3 : 6 14·4			
3 : 4	—	860·8	859·9	+24·0	+ 5·1	2 : 3 : 4 12·4 3 : 4 : 5 17·2	3 : 4 : 6 19·7	3 : 4 : 6 19·7			
3 : 5	—	853·9	853·0	+30·9	+12·0	2 : 3 : 5 14·6 3 : 4 : 5 10·3	—	2 : 3 : 5 14·6			
Trinitrotoluene—											
α (2 : 4 : 6)	227	822·5	820·7	+29·1	+10·2	—	—	—	4·88	0·08	114
β (2 : 3 : 4)	—	834·7	832·9	+16·9	− 2·0	—	—	—	—	0·73	92
γ (3 : 4 : 6)	—	827·4	825·6	+24·2	+ 5·3	—	—	—	—	1·40	102
δ (3 : 4 : 5)	—	829·9	828·1	+21·7	+ 2·8	—	—	—	—	1·60	95
ε (2 : 3 : 5)	—	825·6	823·8	+26·0	+ 7·1	—	—	—	—	0·40	101
ζ (2 : 3 : 6)	—	827·1	825·3	+24·5	+ 5·6	—	—	—	—	—	—

The heats of formation (carbon amorphous) (Table VII) of this series, range from 15 to 31 kgrm.-cal./mol. and in general rise to a maximum for the third (dinitro) member of the series. This maximum in the heats of formation is shown by fig. 1. Thus the course of the nitration of α-trinitrotoluene is traced from toluene, through orthonitrotoluene and 2:4 and 2:6-dinitrotoluenes. The heats of formation of the liquid substances as set forth in this curve are corrected to the solid form. Toluene is assumed to have heat of crystallisation 2·4 cal./mol. and the mononitrotoluenes 4·3 cal./mol.

Some regularities have been observed in the heats of formation of members of the toluene series. Thus there appears to be a definite increase in the internal energy (lower value for heat of formation) associated with groups in the ortho position, and especially when the nitro-groups are ortho to one another. A phenomenon which may be associated with this has been observed by Le Bas* who found in the case of the mononitrotoluenes that the greatest decrease in molecular volume was with ortho-substitution. For the present data for heat of formation the following expression approximately holds:—

Heat of formation of di- and tri-nitrotoluenes $= 30{\cdot}8 - 1{\cdot}3\,(n \times 4m\epsilon)$ where n is the number of nitro-groups ortho to a methyl-group, and m is the number of nitro-groups ortho to a nitro-group.

The extent to which this obtains is shown in the following Table:—

Table VIIA.

Substance. (solid).	Heat of formation.		Substance (solid).	Heat of formation.	
	Calculated.	Experimental.		Calculated.	Experimental.
DNT—			TNT—		
2 : 4	29·5	31·1	α	28·2	29·1
2 : 3	24·3	24·3	β	19·1	16·9
2 : 5	29·5	28·7	γ	24·3	24·2
2 : 6	28·2	29·6	δ	20·4	21·7
3 : 4	25·6	24·0	ε	24·3	26·0
3 : 5	30·8	30·9	ζ	23·0	24·5

Heats of nitration of the toluene series are calculated from heats of formation and are tabulated in Table VII. The heats of nitration of the liquid isomers are uncorrected for change of state, and the amounts of the heat evolved at each stage of the nitration from toluene to the trinitro-compounds are given for all possible entries of the NO_2 group into the ring. The heat of nitration ranges from +33·7 kgrm.-cal./mol. for toluene to *p*-nitrotoluene to 5·3 kgrm.-cal./mol. for 2:4 dinitrotoluene to *β*-trinitrotoluene.

---

* 'The Molecular Volumes of Liquid Compounds,' p. 220.

This Table is made clearer by reference to fig. 3. The formulæ of the isomers are given and the methods of derivation of the substances are indicated. The numbers show the amount of heat liberated during the nitration when one NO_2 group is introduced in the ortho-, meta and para-positions to existing nitro-groups. The insertion of a nitro-group into the *o*-, *m*- and *p*-positions relative to existing nitro-groups, and to the methyl group is indicated in the diagram.

(b) *Benzene Series.*—The results are 4–5 kgrm.-cal./mol. higher than those of Swarts but agree with those of Berthelot (Table I).

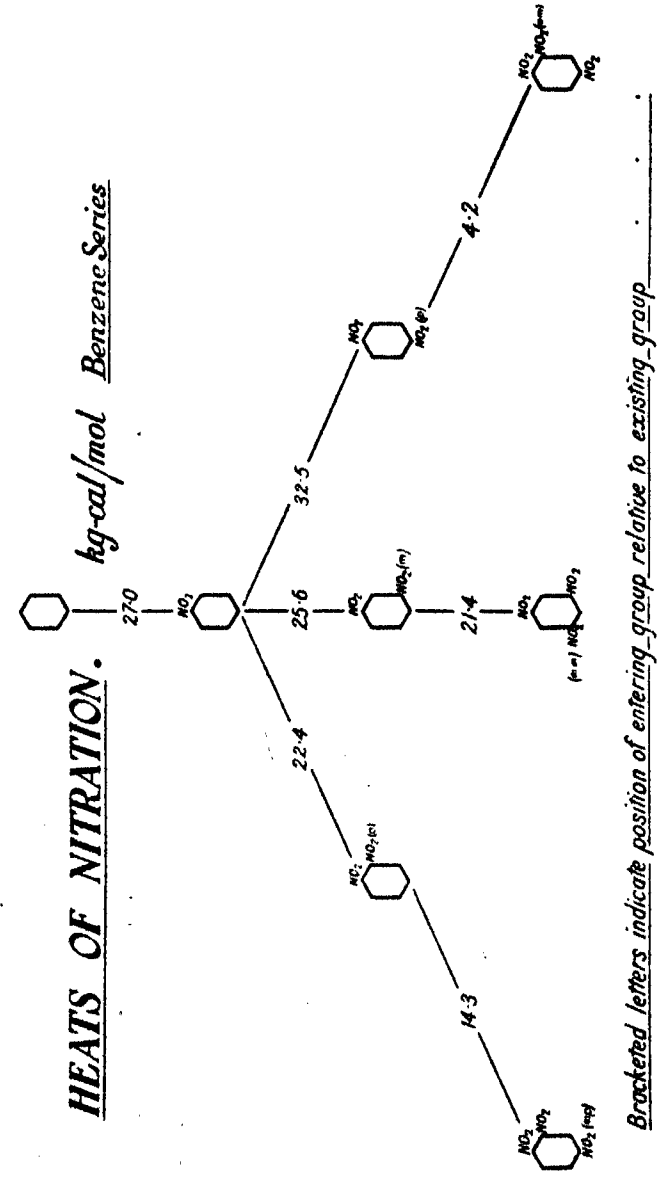
Table VIII.—Heats of Combustion—Benzene Series.

Substance.	Mol. wt.	Cal./grm. in oxygen uncorrected for nitric acid.		Kgrm.-cal./mol. in oxygen uncorrected for nitric acid.	Kgrm.-cal./mol. corrected for nitric acid.
Benzene	78	—		—	780·4 (Richard's value).
Nitrobenzene	123	6022 6036 6020 6028 6018	6025	741·1	739·9
o-Dinitrobenzene	168	4195 4191 4207	4198	705·2	703·8
m-Dinitrobenzene	—	4181 4186 4173	4180	702·3	700·6
p-Dinitrobenzene	—	4140 4130 4144 4137	4138	695·2	693·7
1 : 3 : 5 Trinitrobenzene	213	3126 3133 3140	3133	667·3	665·6
1 : 2 : 4 Trinitrobenzene	—	3177 3172 3186 3174	3177	676·9	675·9

The heats of formation (Table IX) are lower than for the toluenes and do not give a maximum such as that obtained for the nitrotoluenes; the values rise from +5·6 for benzene to +26·1 for *p*-dinitrobenzene. The nature of the rise to the symmetrical trinitrobenzene is shown in fig. 1. Of the two isomeric trinitrobenzenes the symmetrical has the greater heat of formation and also possesses greater stability and has a higher figure of insensitiveness than the unsymmetrical compound (see below). This agrees with the behaviour of the trinitrotoluenes.

Table VIIIA.

Explosive.	Heat of formation (kgrm.-cal./mol.).	Figure of insensitiveness (picric acid = 100).	Gas evolution (c.c. in 40 hours) at 120° C.
1 : 3 : 5 trinitrobenzene.........	21·1	106	Nil
1 : 2 : 4 trinitrobenzene.........	10·8	103	0·66



Bracketed letters indicate position of entering group relative to existing group

FIG. 4.

Table IX.—Heats of Combustion, Formation, Nitration and Crystallisation—Benzene Series.

Substance.	Mol. wt.	Heat of combustion, kgrm.-cal./mol.		Heat of formation, kgrm.-cal./mol.		Heat of nitration, kgrm.-cal./mol.	Heat of crystallisation, kgrm.-cal./mol.
		Const. vol.	Const. press.	C. amorph.	C. diamond.		
Benzene	78	780·4	781·3	+ 5·6	−10·6	+27·0	2·35
Nitrobenzene	123	739·9	739·7	+13·1	− 3·1	+25·6	2·78
o-Dinitrobenzene	168	703·8	702·6	+16·0	− 0·2		
m-Dinitrobenzene	168	700·6	699·4	+19·2	+ 3·0	+21·4	4·87
p-Dinitrobenzene	168	693·7	692·5	+26·1	+ 9·9		
1 : 3 : 5-Trinitrobenzene	213	665·6	663·4	+21·1	+ 4·9		
1 : 2 : 4-Trinitrobenzene	213	675·9	673·7	+10·8	− 5·4		

Table XI.—Heats of Combustion, Formation, Nitration and Crystallisation—Phenol Series.

Substance.	Mol. wt.	Heat of combustion, kgrm.-cal./mol.		Heat of formation, kgrm.-cal./mol.		Heat of nitration, kgrm.-cal./mol.	Heat of crystallisation, kgrm.-cal./mol.
		Const. vol.	Const. press.	C. amorph.	C. diamond.		
Phenol	94	736·9	737·5	+49·4	+33·2	+35·0	2·34
o-Nitrophenol	139	689·9	689·4	+63·4 } 64·9	+47·3 } 48·7		
p-Nitrophenol	139	687·0	686·5	+66·3	+50·1	+24·5	4·30
2 : 4-Dinitrophenol	184	650·2	648·7	+69·9	+53·7	+12·9	
Picric acid	229	623·7	621·2	+63·3	+47·1		

The curve of heats of nitration is not so steep as for the phenol and methylaniline series, and this appears to be connected with the difficulty with which the nitro-group enters the molecule. The slope corresponding to the conversion of dinitrobenzene into trinitrobenzene is very slight, and it is well known that this step in the nitration proceeds only with great difficulty.

The variations in the heats of nitration of the isomers in the benzene series is indicated in fig. 4, where the same method is adopted as in fig. 3. The relative order of the heats of nitration into the ortho-, meta- and para-positions is similar to that found for the toluene series, ortho-substitution taking place with evolution of the least amount, and the para-substitution with the greatest amount of heat.

(c) *Phenol Series.*—The heats of combustion of the phenols are very low, owing to the presence of a hydroxyl group. The values for *o*- and *p*-nitrophenol are of the same order as those given in Table I. It should be noted, however, that the *o*-nitro-compound has the higher heat of combustion.

Table X.—Heats of Combustion—Phenol Series.

Substance.	Mol. wt.	Cal./grm. in oxygen.	Kgrm.-cal./mol. uncorrected.	Kgrm.-cal./mol. corrected.
Phenol	94	—	—	736 ·9 (Berthelot's value).
o-Nitrophenol	139	4962 4972 } 4972 4981	691 ·1	689 ·9 } 688 ·5
p-Nitrophenol	139	4950 4950 } 4951 4953	688 ·2	687 ·0
2 : 4 Dinitrophenol	184	3540 3546 } 3543	651 ·9	650 ·2
Picric acid..........	229	2728 2727 } 2731 2743	625 ·4	628 ·7

(d) *Methylaniline Series.*—The heats of formation rise to a maximum for the third member, dinitromethylaniline, and then rapidly decrease, until, for tetryl, a heat of formation + 7·5 is obtained.

Table XII.—Heats of Combustion—Methylaniline Series.

Substance.	Mol. wt.	Cal./grm. in oxygen.	Kgrm.-cal./mol. uncorrected for nitric acid.	Kgrm.-cal./mol. corrected for nitric acid.
Methylaniline	107	9127 9114 } 9119 9116	975·7	974·4
p-Nitromethylaniline	152	6090 6093 } 6090 6088	925·8	924·3
2 : 4-Dinitromethylaniline	197	4504 4510 } 4507 4506	887·9	885·5
2 : 4 : 6-Trinitromethylaniline	242	3562 3563 } 3564 3566	862·5	859·9
Tetryl	287	2950 2951 2955 } 2955 2957 2964	848·1	845·3
Methyl Tetryl	301	3375 } 3371 3366	1014·7	1012·1

Table XIII.—Heats of Combustion, Formation and Nitration—Methylaniline Series.

Substance.	Mol. wt.	Heat of combustion, kgrm.-cal./mol.		Heat of formation, kgrm.-cal./mol.		Heat of nitration, kgrm.-cal./mol.
		Const. vol.	Const. press.	C. amorph.	C. diamond.	
Methylaniline	107	974·4	975·4	+11·0	− 7·9	+36·4
p-Nitromethylaniline	152	924·3	924·3	+27·9	+ 9·0	+25·2
2 : 4-Dinitromethylaniline	197	885·5	884·5	+33·5	+14·7	+11·9
2 : 4 : 6-Trinitromethylaniline	242	859·9	857·9	+26·0	+ 7·1	+ 1·0
Tetryl	287	854·3	842·3	+ 7·5	−11·4	
Methyl tetryl	301	1012·1	1009·3	+ 5·8	−15·8	

Table XIV.—Heat of Entry of Nitro-Group (into Solid Substances).

Substance.	*o*-position.	*m*-position.	*p*-position.
Toluene	+7·7	+11·8	+11·8
Nitrotoluene—			
2 (ortho)	2 : 3−0·9	2 : 4+5·9 2 : 6+4·4	2 : 5+3·5
3 (meta)	2 : 3−5·0 3 : 4−5·8	3 : 5+1·6	2 : 5−0·6
4 (para)	3 : 4−5·8	2 : 4+1·8	
Dinitrotoluene—			
2 : 3	2 : 3 : 4− 7·4	2 : 3 : 5+1·7	2 : 3 : 5+1·7 2 : 3 : 6+0·2
2 : 4	3 : 4 : 6− 6·9 2 : 3 : 4−14·2	2 : 4 : 6−2·0	3 : 4 : 6−6·9
2 : 5	2 : 3 : 5− 2·7 3 : 4 : 6− 4·5 2 : 3 : 6− 4·2	2 : 3 : 5−2·7 3 : 4 : 6−4·5 2 : 3 : 6−4·2	
2 : 6	2 : 3 : 6− 5·1	2 : 4 : 6−0·5	2 : 3 : 6−5·1
3 : 4	2 : 3 : 4− 7·1 3 : 4 : 5− 2·8	3 : 4 : 6+0·2	3 : 4 : 6+0·2
3 : 5	2 : 3 : 5− 4·9 3 : 4 : 5− 9·2	—	2 : 3 : 5−4·9

Conclusions.

The heats of formation of the nitro-compounds of the series investigated are markedly influenced by the position of the nitro-groups, and there is no regular change in these values similar to that observed in the introduction of the CH_2 group into aliphatic compounds; no empirical equation can be devised showing the relationships between the heats of formation of all aromatic nitro-compounds. From an examination of figs. 1 and 2, however, it is seen that certain regularities exist between the heats of formation and nitration of the benzene, toluene, phenol and methylaniline series. The heats of combustion in all of the series investigated decrease with increase in the number of nitro-groups, but the heats of formation tend to a maximum for the second or third member of the series, and then diminish. An exception is seen in the benzene series, but here it is probable that the maximum is shifted to the fourth member of the series, and that the same general relationship would hold if the more highly nitrated benzenes could be obtained.

It has been found in the toluene series that there is some regularity in the heats of formation of the di- and tri-nitroderivatives, and an expression is given on p. 227 which takes into account the lowering of this constant by nitro-groups, ortho to methyl, and to one another.

The introduction of the methyl group into benzene, to form toluene, modifies only slightly the shape of the curves showing the heat of formation of the nitro-compounds, but the introduction of a hydroxyl- or methylamino-group has a marked effect. The resemblance between the curves for the phenol and methylaniline series is striking; the effect of the hydroxyl-group on the energy changes involved in nitration is thus very similar to that of the methylamino group, a result scarcely to be expected. The heats of nitration curves for these substances are straight lines. On the other hand, there is no simple additive relation for the entry of nitro-groups into the benzene and toluene series. The differences between the benzenes and toluenes on the one hand, and the phenols and methylanilines on the other, is apparently connected with the relative ease of nitration of these substances; the nitration of the former, especially to the highly nitrated members, takes place with difficulty, whereas the latter are easily nitrated.

In this connection also it may be mentioned that it is possible to nitrate phenol to the tetranitro-derivative, and methylaniline to the tetranitrophenylmethylnitramine and pentanitrophenylmethylnitramine, while only the trinitro-derivatives of benzene and toluene have been obtained.

From the results of this study of the nitro-compounds of the benzene, toluene and phenol series, it is clear that the nearer the nitro-groups are to one another in the benzene ring the greater is the strain in the molecule and the lower the heat of formation. Thus the *o*-mononitro-compounds of these series have the smallest heat of formation; the differences between the ortho- and para- derivatives are, however, greatest for benzene and least for the phenol series. In the case of di- and tri- nitro-derivatives also, the proximity of the groups has the same effect.

The sensitiveness of nitro-compounds to impact and their comparative instability run in line with their heats of formation; thus β-trinitrotoluene, which has the lowest heat of formation, is the most sensitive to impact. The stability of the unsymmetrical trinitrotoluenes and benzenes is less than that of the symmetrical forms.

Certain conclusions can now be made with regard to the part played by the heat of formation in determining the magnitude of the heat of detonation of a high explosive. The large heat of formation of the nitro-phenols is a disadvantage from the point of view of their employment as explosives, though this is to some extent balanced by their comparatively high oxygen-content. Thus, although trinitrotoluene contains 29 per cent. less of the oxygen necessary for complete combustion than does picric acid, yet its heat of detonation is similar, namely, 924 cal. per gramme (water gaseous), compared with 914 cal.* per gramme (water gaseous). This is to be expected from the lower heat of formation of the trinitrotoluene (128 calories as compared with 277 calories per gramme for picric acid). In this example the introduction of the hydroxyl group into benzene to form phenol is accompanied by the absorption of more energy than that of the methyl group to form toluene, and this is reflected in the explosive phenomena.

---

* Robertson, 'Trans. Chem. Soc.,' vol. 119, p. 8 (1921).

From the figures that are available, the fifth member of a series, *i.e.*, a tetranitro-compound, will be seen to have a very low heat of formation. The effectiveness of tetryl as an explosive is very largely dependent on its being the fifth member of the methylaniline series and thus possessing a low heat of formation. The entry of the fourth nitro-group into the side-chain does not affect this uniformity.

Of the isomeric trinitrotoluenes, β-trinitrotoluene has been shown to have the lowest heat of formation, *i.e.*, to contain the lowest internal energy per gramme, and α-trinitrotoluene the highest heat of formation, *i.e.*, to contain the least internal energy per gramme. If on detonation these two substances gave the same products of decomposition, β-trinitrotoluene would have an advantage over α-trinitrotoluene of 54 cal. per gramme. For the same reason, the unsymmetrical trinitrobenzene should be a more powerful explosive than the symmetrical body. In general, the greater the number of adjacent groups in a molecule of a high explosive, the greater will be the heat of detonation, but there are often practical limitations to this choice on account of the greater reactivity and lower stability of the less symmetrical nitro-compounds.

We wish to tender our thanks to the Director of Artillery for permission to publish this work, and to Sir Robert Robertson for the interest he has taken in the investigation.

Researches upon Brown Coals and Lignites. Part I.—*Heat Treatment at Temperatures below 400° C. as a Possible Method for Enhancing their Fuel Values.*

By William A. Bone, D.Sc., F.R.S.

(Received February 10, 1921.)

Introduction.

Coals may be conveniently divided into four principal genera, namely (A) Sub-bituminous, including brown coals and lignites, (B) Bituminous, (C) Semi-bituminous, and (D) Anthracitic, of which the first is, generally speaking, of more recent origin than the other three. According to the Report issued by the International Geological Congress in 1913, of the world's estimated total coal reserves (7,397,553 million tons), about 40·5 per cent. are sub-bituminous and lignitic, 52·75 are bituminous, and the remaining 6·75 per cent. are anthracitic. It is evident, therefore, that a large proportion of the world's coal reserves are of the immature sub-bituminous type whose economic values as fuels have hitherto been regarded as much inferior to those of the geologically older bituminous and anthracitic groups.

The world-wide economic dislocations caused by the War, and notably the great increase in the cost of winning and transporting the more valuable bituminous and anthracitic coals, have so profoundly affected the fuel situation generally, that the relative standing of the sub-bituminous brown coals and lignites has been of late much enhanced. Indeed, it may be truly said to-day that in many countries, including several of our Overseas Dominions, the question of how best to utilise their potential lignite coal resources has become one of first-rate importance. Already in Australia, Canada, and the United States, the matter is being vigorously investigated under Government auspices or direction as one of pressing economic necessity. It may, therefore, be assumed that if scientific research could disclose how the inferior brown coals and lignites may be "up-graded" by a simple process into something approximating to the more valuable black coals, it might not only add to our knowledge of coal formation in past ages, but also would powerfully aid the solution of a present-day economic problem of world-wide importance. In this paper will be described the results of some experiments which have been recently carried out with this end in view.

Lignites and brown coals comprise a great variety of principally tertiary coals which may be regarded as intermediate between peat and bituminous coals. In their natural state they vary considerably in character and chemical composition according to the degree of transformation of the original vegetable *débris* (chiefly lignin) from which they were derived. From this point of view, further investigation may ultimately show that they include as many sub-groups as are known to exist among the geologically older bituminous (*i.e.*, black) coals. Unfortunately, however, our present chemical knowledge of the brown coals is too limited to admit of anything better than some such provisional classification of them, according to their external appearances, as the following:—

(A) Woody or Fibrous Brown Coals, such as retain a distinct ligneous structure;

(B) Amorphous or Earthy Brown Coals, which are devoid of any organic structure;

(C) Common or Brown Lignites, which are dark brown in colour and have a distinct laminated structure and fracture; and

(D) Black Lignites, which are nearly black in colour and have a conchoidal fracture.

It is also possible that such a classification gives some index of the extent to which the "coal formation" has proceeded.

These sub-bituminous coals are further characterised as a genus by the following common features, which, taken together, differentiate them from the older bituminous coals, namely, in that (i) they naturally contain a large amount of water (*e.g.*, always more than 10, and sometimes even as much as 50, per cent.); (ii) on losing the same by "air-drying" they usually disintegrate, either crumbling to powder or developing well-marked laminar cracks; (iii) they are absolutely devoid of any coking properties; and (iv) in the "dry ashless" state they contain much more oxygen and less carbon than do any of the bituminous coals. Thus a "dry ashless" lignite will usually contain less than 70 per cent. of carbon and more than 20 per cent. of oxygen.

Owing to their low carbon and high oxygen contents, sub-bituminous coals have, even in the dry state, much lower calorific values than any of the older types; and when their high water-contents in the raw state are considered, together with the fact that they usually disintegrate on being dried by any of the ordinary methods, it will be readily understood that the problem of their transportation for any long distance has always presented formidable difficulties.

In the raw state they are at best a low-grade fuel, burning with a very long and smoky flame. And, owing to so much heat being required to evaporate their natural water-content, they do not give a very hot or radiating fire, which in boiler practice is so essential to rapid heat transmission and efficient working. Against their obvious inferiority as fuels to bituminous coals may be set the consideration that some of them are capable of yielding on carbonisation at low temperatures a variety of valuable by-products including volatile hydrocarbons, fuel and lubricating oils, solid paraffins, etc. These latter possibilities will be more fully dealt with in a subsequent paper, but meanwhile attention may be drawn to them.

As already mentioned, most of the brown coals and lignites are geologically of tertiary age, although some of them are undoubtedly cretaceous. Indeed, it has been proposed to discriminate between what have been termed true brown coals or lignites, which are chiefly tertiary, and sub-bituminous coals proper, which are chiefly cretaceous. This is, however, more of a geological than a chemical question, which, lying as it does outside the ambit of the author's purpose, does not immediately concern him.

Tertiary brown coals and lignites are chiefly found (i) in what may be termed the "Pacific borderlands," *e.g.*, in North-west America, Japan, Australia, New Zealand, and the East Indies; (ii) in regions adjacent to, or in continuation with, the Gulf of Mexico; and (iii) in regions north of the Mediterranean, including the Central European Plain.

Economically speaking, the most important of the known tertiary lignite areas in the world are probably those of Victoria (Australia), Saskatchewan (Canada), North Dakota (U.S.A.), and Germany, although valuable deposits occur, and in some cases are being worked, in Japan, India, Burma, Malaya, and New Zealand, as well as other parts of our Overseas Dominions. Cretaceous sub-bituminous coals chiefly occur in the western parts of Canada (Alberta) and the United States.

A. *An Experimental Investigation of Morwell Brown Coal.*

The present research originated in a request, made to the author in October, 1916, by the Agent-General for the Government of Victoria (Australia), for certain assistance in connection with examining the nature of the tars and oils obtained on carbonising at a low temperature the celebrated Morwell brown coal deposits in the Gippsland district, about 97 miles from Melbourne. These Victorian brown-coal deposits are of tertiary origin, and, according to the Report of the Government Advisory Committee, they consist of "a matrix of earthy brown coal with sporadic inclusions of lignite. . . . The matrix consists of pollen grains, spore cases, and decomposed vegetable matter. . . . The coal varies in colour between yellowish brown and black, but it always pulverises to brown powder."* The Morwell deposits, which lie quite near the surface, are of exceptional thickness, and represent an enormous store of potential energy whose economic importance to the State of Victoria can scarcely be overrated. Thus, in the Report of the International Geological Congress upon the Coal Resources of the World, it is stated that "the thickness of the Gippsland deposits is shown in a bore-hole put down near Morwell which disclosed no fewer than seven beds of brown coal. Three of these proved to be of phenomenal thickness, without parallel elsewhere."* A summary of the seams cut was as follows:—

---

* 'Report of Victorian Advisory Committee on Brown Coal,' 1917.

Depth from surface to top of seam.		Thickness of coal.	
ft.	ins.	ft.	ins.
66	2	29	8
167	6	25	8
202	2	23	0
251	6	227	10
482	4	265	6
818	4	166	1
995	2	43	8
Total coal		781	5

Unfortunately, however, these Morwell brown coals usually contain in the raw state between 45 and 55 per cent. of water, and therefore it is difficult to burn them efficiently under boilers. Nevertheless, in the year 1917, the Advisory Committee appointed by the Victorian Government to investigate the possibilities of generating electric power from them on a large scale, reported† that, notwithstanding their low-grade character, it would be more economical to generate electric power for the City of Melbourne from them than from New South Wales coals. About that time also it was officially estimated that the cost of producing raw Morwell coal at the mines would not exceed 2*s.* 9*d.* per ton. The estimated cost of delivering Morwell coal in Melbourne was then 7*s.* 8*d.* per ton, against about 20*s.* for black coal imported from Newcastle (New South Wales). And as the result of developments that are now taking place, it is anticipated that, in a few years hence, not only will the City of Melbourne derive the whole of its electric power from Morwell coal, but that the Victorian State Railways will also be worked electrically by energy from the same deposits.

A representative sample of this Morwell coal was, in 1917, forwarded to the author for scientific investigation from the Victorian Ministry of Mines, through the courtesy of Sir Peter MacBride, the Agent-General for Victoria in London. It was consigned in six sealed tin cases, and was found to be a typical earthy brown coal. Its moisture-content, as it reached his laboratory, was 48 per cent. This water was all rapidly eliminated when the coal was maintained either (*a*) at 105° C. in a current of nitrogen, or (*b*) at the room temperature in a vacuum over strong sulphuric acid. The dry coal yielded 55·3 per cent. of volatile matter when carbonised at 900° C. out of contact with air, and its ultimate composition was found on analysis to be as follows:—

---

* 'Coal Resources of the World,' vol. 1, p. 42.

† 'State of Victoria Report of the Advisory Committee on Brown Coal,' Melbourne, 1917.

	Dry coal.	Dry ashless coal.
	per cent.	per cent.
Carbon	62·50	65·10
Hydrogen	4·85	5·00
Nitrogen	0·45	0·45
Sulphur	0·20	0·20
Oxygen	28·00	29·25
Ash	4·00	—
	100·00	100·00

Gross calorific value of the dry coal = 5600 K.C.Us. per kilogramme (or 10,080 B.T.Us. per lb.).

Heat Treatment of the Dry Coal.—On investigating the behaviour of the thoroughly dried coal when further heated, out of contact with air, in a special form of apparatus which permitted accurate control of temperature and the measurement of the amounts of either liquid or gaseous products evolved, the following significant observations were made, which subsequent research has proved to be characteristic of brown coals and lignites generally. It was found that, beginning at as low a temperature as 130° C., but principally between 250° and 375°, a chemical change occurred, which was characterised by the elimination of both steam and carbon dioxide from the coal substance, without any appearance of condensable oils or more than a quite negligible amount of gaseous hydrocarbons. Thus, in a typical experiment, when the temperature of 100 parts by weight of the dry coal was slowly raised to, and then maintained at, 375° C. until no further change occurred (the whole duration of the operation averaging about two hours), there were eliminated 5·5 parts of water and 6·6 parts by weight of gas (= 1427 cubic feet of gas per ton of dry coal), which latter contained:—

$H_2S = 1{\cdot}5$, $CO_2 = 88{\cdot}5$, $CO = 4{\cdot}1$, $CH_4 = 1{\cdot}1$, and $N_2 = 4{\cdot}8$ per cent.

The reason for not continuing the heating beyond 375° C. was that such degree marked the limit to which the treatment could be pushed without the coal losing either hydrocarbon oils or gases. And—inasmuch as subsequent investigation showed that for any particular lignite there is a similar temperature limit (usually between 300° and 400° C.) up to which the dry "coal substance" in like manner loses large quantities of both carbon dioxide and steam as the result of a chemical condensation affecting mainly (if not entirely) its cellulosic or humic constituents, without either appreciably changing its hydrocarbon-yielding constituents or materially diminishing its potential energy as a fuel,—it is proposed to term such temperature the "practicable upgrading limit" for that particular lignite. This phrase is suggested because the temperature in question marks the practicable limit beyond which the concentration of potential energy in the dried fuel, consequent upon the said chemical condensation, cannot usefully be pushed without incurring appreciable thermal loss.

How far such concentration of potential heat units is practicable in the case of the Morwell coal investigated may be deduced from the following considerations, namely:—

(1) That the 87·6 parts of "residual coal" obtained after treating 100 parts by weight of the dry coal in the foregoing experiments was found on analysis to contain:—

	Per cent.
Carbon	68·5
Hydrogen	4·9
Nitrogen	0·5
Sulphur	0·3
Oxygen	21·2
Ash	4·6
	100·0

(2) That its gross calorific value was:—

6,360 K.C.Us. per kilogramme (or 11,450 B.T.Us. per pound);

(3) And, therefore, that the chemical and thermal balances for the experiments were somewhat as follows:—

	100 parts of the original dry coal containing	yielded	*87·6 parts of residual coal containing*	Loss.
	parts.		parts.	parts.
Carbon	62·50		60·00	2·50
Hydrogen	4·85		4·30	0·55
Oxygen	28·00		18·55	9·45
K.C.Us.	560,000		558,000	2,000

It will thus be seen that the original dry coal had lost altogether 13·4 per cent. of its weight (principally as carbon dioxide and steam), but had retained practically the whole of its potential heating power. The loss in question had amounted to approximately one-third of the oxygen, one-tenth of the hydrogen, but only one-thirtieth of the carbon originally present in the coal substance. Leaving the bearing of such a result upon the chemistry of coal formation to be discussed at some future time, some of its possible economic applications may now be indicated.

Some Economic Considerations.—It being obvious from the foregoing experiments that the brown coal in question could be considerably "up-graded" by a simple heat treatment at temperatures well below 400° C. in the manner described, further investigations were made with a view to determining whether the heat required first to dry and then to "up-grade" the raw coal could be supplied in a large power station or boiler installation at the expense of a portion of the sensible heat in the products of combustion from the treated fuel, assuming them to leave the boiler system at a temperature of not less than 400° C. These investigations have shown that such combustion products do carry away from the boiler several times more heat than would suffice for the purpose in view. Moreover, laboratory experiments have indicated that it would not be necessary, for the completion of the "up-grading" process, to maintain the dry coal at 375° C. for more than about forty-five minutes, and also that the aforesaid products of combustion could be safely employed for the purpose provided that they contain more than 10 per cent. of carbon dioxide. Large-scale trials of the scheme are accordingly now in contemplation, with a view to testing its applicability to the Morwell deposits.

The advantages accruing from a successful issue of such trials would be (*a*) that the raw fuel would be dried and "up-graded" *in situ* by means of heat which would otherwise be lost up the chimney, or, in other words, that the incoming raw coal would act as an "economiser" in the boiler system; and (*b*) that the "up-graded" fuel would subsequently give a much hotter and more radiant fire in the boiler furnace, thereby increasing the steam output and thermal efficiency of the system.

B. *Investigation of other Typical Lignites.*

As it seemed important to ascertain whether (and if so, then how far) the behaviour of Morwell brown coal under the aforesaid heat treatment is characteristic of sub-bituminous coals as a class, steps were taken to obtain representative samples of lignites of widely different origins. The author was fortunate in securing several such typical consignments, from which the following were selected as best answering the purpose in view:—

(1) A series of Italian woody lignites (Class A) from the San Giovanni Valdarno District.

(2) Two Italian lignites (Class C) from Umbria.

(3) A series of four Canadian lignites (Class C/D) from the Province of Saskatchewan.

(4) A Burmese lignite (Class D).

Details of the investigation of all these samples will now be given.

(1) *Italian Lignites from the Valdarno District.*—There are several lignite deposits in the northern parts of Italy (chiefly in Tuscany and Umbria) of which those of San Giovanni Valdarno are economically the most important. In pre-war days the Italian production of lignite amounted to about 550,000 tons per annum, of which about 75 per cent. were of Tuscan, and about 20 per cent. of Umbrian origin. The principal Italian deposits occur in the mountain chain separating the valley of the Arno from that of Chianti, and are commonly known as the Valdarno deposits.

According to the Report issued by the International Geological Congress in 1913, these Valdarno deposits occur in the lower Pliocene measures, and in the raw state they usually contain between 40 and 50 per cent. of water. They have a very woody appearance and fibrous structure, so much so that they can usually be sawn into pieces without crumbling. In colour they are dark brown. The author was fortunate in securing through the Italian Military Mission in London one representative sample, and subsequently he obtained some others privately.

As received at his laboratory, they contained between 25 and 40 per cent. of water; the volatile matter in the "dry ashless" coals varied between 55 and 60 per cent. An ultimate analysis of one of the dried samples (No. 1) was made, with the following results, from which it will be seen that (as might be expected from its "woody" character) the coal substance contained decidedly less carbon but more oxygen than the Morwell brown coal.

On investigating the behaviour of three different consignments of Valdarno lignites when submitted in a dry state to the heat treatment in question, it was found that the highest temperature to which they could be raised without losing hydrocarbon or volatile oily vapours varied between 285° and 310° C., instead of the 375° previously employed in the case of the Morwell coal. Possibly this difference may be attributed to the more woody (*i.e.*, resinous) character of the Valdarno coals. In other respects their behaviour was very similar to that of the Morwell coal, namely, that up to the experimental temperature limit in question, water and carbon dioxide were freely eliminated from the dried coal substance without any material loss in potential calorific value. It should, however, be noted that the gases eliminated from the coal

Valdarno No. 1.

	Dry coal.	Dry ashless coal substance.
	per cent.	per cent.
Carbon	54·90	60·30
Hydrogen	4·90	5·40
Nitrogen	0·60	0·65
Sulphur	1·20	1·30
Oxygen	29·50	32·35
Ash	8·90	—
	100·00	100·00
Percentage volatile matter yielded at 900° C.	52·6	57·8

Gross calorific value of the dry coal = 5390 K.C.Us. per kilogramme (=9710 B.T.Us. per lb.).

(1) Results of Experiments upon Italian Lignites.

	Valdarno (Tuscany).			Umbria.	
	1.	2.	3.	Fontivecchi.	Poggio Mirteto.
Per cent. moisture in raw coal	39·1	25·4	27·4	21·9	28·6
Data relating to dry coal—					
Percentage volatiles yielded at 900° C.	52·6	52·3	53·5	48·5	53·5
Percentage ash	8·9	6·0	5·7	16·3	9·6
Gross calorific value (K.C.Us. per kgrm.)	5390	5490	5690	4595	5630
Experimental temperature limit (C°.)	300°	310°	285°	320°	300°
Results of heat treatment of the dry coal—					
Percentage loss in weight	8·0	9·7	7·5	10·4	9·9
	Gas 4·0 H_2O 4·0	Gas 4·0 H_2O 5·7	Gas 2·8 H_2O 4·7	Gas 4·7 H_2O 5·7	Gas 3·2 H_2O 6·7
Cubic feet of gas at N.P.T. per ton	808	786	550	908	624
Percentage composition of the gas—					
CO_2	74·0	85·2	79·6	83·7	82·9
CO	21·0	10·5	12·7	10·9	6·6
C_2H_6	1·3	0·3	2·7	1·0	1·5
N_2	3·7	4·0	5·0	4·4	9·0
Gross calorific value of residual fuel (K.C.Us. per kgrm.)	5830	5840	6010	5060	6020
K.C.Us. in residue from 1 kgrm. dry coal	5365	5265	5505	4425	5445

during the heat treatment contained appreciably more carbon monoxide than did those previously obtained in the experiments with Morwell coal. The general character of the experimental results may be judged from Table (1) on p. 244.

The gross calorific value of the carbonaceous residue obtained from the No. 1 sample was:—

5830 K.C.Us. per kilogramme (= 10,500 B.T.Us. per lb.),

and it had the following ultimate composition:—

	As obtained.	On "ashless" residue.
	per cent.	per cent.
Carbon	58·9	65·2
Hydrogen	4·7	5·2
Nitrogen and sulphur	2·0	2·2
Oxygen	24·7	27·4
Ash	9·7	—
	100·0	100·0

The following balance relating to the heat treatment of the consignment in question has been deduced from the experimental data.

	100 parts of the original dry coal containing	yielded	92 parts of residual coal containing	Loss.
	parts.		parts.	parts.
Carbon	54·9		54·2	0·7
Hydrogen	4·9		4·3	0·6
Oxygen	29·5		22·7	6·8
K.C.Us.	539,000		536,500	2,500

It will thus be seen that, as the result of the heat treatment of the particular sample, practically 99 per cent. of the original potential energy of the coal had been concentrated in the ninety-two parts of the residual fuel. From the thermal point of view, the results with the other Valdarno samples were perhaps not quite so favourable; in these cases there was, on the average, a concentration of about 97 per cent. of the original potential energy of the dried coal in the ninety parts of the residual fuel obtained.

(2) *Laminated Italian Lignites from Umbria.*—Two consignments of Umbrian lignites from Fontivecchi and Poggio Mirteto, respectively, were next investigated. They were quite different in both colour and structural

character from the Valdarno samples, having a dark greyish-brown appearance, and laminated structure with "slaty" cleavage. Umbrian lignites are usually considered to be of Pliocene origin, and, in pre-war days, between 100,000 and 150,000 tons per annum were produced for fuel purposes. As received at the author's laboratory, the raw coals contained between 20 and 30 per cent. of water.

The following data for the ultimate chemical composition of the dry ashless Fontivecchi sample show that it approximated closely to that of the Valdarno coals:—

	Dry coal.	Dry ashless coal substance.
	per cent.	per cent.
Carbon	51·3	61·3
Hydrogen	4·2	5·0
Nitrogen and sulphur	1·0	1·2
Oxygen	27·2	32·5
Ash	16·3	—
	100·0	100·0

Gross calorific value = 4595 K.C.Us. per kilogramme (8280 B.T.Us. per lb.).

Further particulars concerning the character of the coals, and their behaviour when subjected to the heat treatment in question, are tabulated on p. 244; it will be seen that the temperature to which they could be raised without losing hydrocarbons was between 300° C. and 320° C. Up to such temperature limit, the dried coals lost on the average about 10 per cent. of their weight (again chiefly as carbon dioxide and water), and about 97 per cent. of their potential calorific value was concentrated in the carbonaceous residue. The ultimate composition of this residue was as follows:—

Composition of the Residue from Fontivecchi Coal.

	As obtained.	On "ashless" residue.
	per cent.	per cent.
Carbon	58·0	71·3
Hydrogen	3·6	4·4
Nitrogen and sulphur	1·0	1·2
Oxygen	18·7	23·1
Ash	18·7	—
	100·0	100·0

Gross calorific value = 5060 K.C.Us. per kilogramme (or 9120 B.T.Us. per lb.).

The balance for the heat treatment was therefore as follows:—

	100 parts of the original dry coal containing	yielded	87·4 parts of residual coal containing	Loss.
	parts.		parts.	parts.
Carbon	51·3		50·7	0·6
Hydrogen	4·2		3·1	1·1
Oxygen	27·2		16·3	10·9
K.C.Us.	459,500		442,500	17,000

(3) *Canadian lignites.*—Through the courtesy of the Lignite Research Board of Canada, four typical samples of lignites from the Province of Saskatchewan were sent to the author for detailed investigation during the summer of 1919. Each of them weighed between 40 lb. and 50 lb., and was shipped over in a sealed tin, marked so as to show its origin, as follows:—

(1) Estevan Coal and Brick Co., Estevan.
(2) Saskatchewan Coal, Brick, and Power Co., Shand.
(3) Bienfait Mine, Bienfait.
(4) Western Dominion Collieries, Taylorton.

All the four samples had a dull brown or slaty black colour, and a conchoidal fracture. They were quite devoid of any "woody" structure, and, although typical tertiary lignites of a character intermediate between Classes C and D, they were certainly more mature coals than any of those dealt with in the preceding section of this paper.

As received at South Kensington, the samples contained between 32 and 38 per cent. of water. When thoroughly dried, they yielded on the average about 44 per cent. of volatile matter at 900° C., or the equivalent of about 48 per cent. on the dry ashless coal substance. Their ultimate compositions and calorific values (for the dry coal) were as follows:—

	1.	2.	3.	4.
	per cent.	per cent.	per cent.	per cent.
Carbon	63·8	62·7	64·5	61·6
Hydrogen	4·2	4·7	4·3	4·2
Sulphur	0·6	0·8	0·7	0·9
Nitrogen	0·6	0·8	0·7	1·0
Oxygen	24·7	22·5	20·5	21·1
Ash	6·1	8·5	9·3	11·2
Calorific values, K.C.Us. per kilogramme	5885	5980	5920	5650

In order, however, that a proper comparison may be made with the corresponding data for the Morwell coal (q.v.), it is necessary to exclude from the calculation the ash contents, and to consider only the following figures for the dry "coal substance" itself:—

	1.	2.	3.	4.
	per cent.	per cent.	per cent.	per cent.
Carbon	68·0	68·5	71·1	69·4
Hydrogen	4·4	5·1	4·7	4·7
Oxygen	26·2	24·6	22·6	23·7

The average potential energy of the Saskatchewan "coal substance" would be 6330 K.C.Us. per kilogramme, as compared with only 5830 for that of the less mature Morwell coal.

The experimental heat treatment of the dry coals, the results of which are set forth in the Table (2) below, showed that, up to a temperature limit varying between 320° C. and 360° C. (but on the average of about 350°), they lost about 8·5 per cent. in weight by the expulsion of steam and oxides

(2) Results of Experiments upon Canadian Lignites.

Description of sample.	1. Estevan.	2. Shand.	3. Bienfait.	4. Taylorton.
Per cent. moisture in raw coal as received	37·8	37·1	35·8	31·8
Data relating to dry coal—				
Percentage volatiles yielded at 900° C.	41·9	48·7	41·5	44·8
Percentage ash	6·1	8·5	9·2	11·2
Gross calorific value (K.C.Us. per kgrm.)	5885	5980	5920	5650
Experimental temperature limit (C.)	330°	360°	320°	340°
Results of heat treatment of the dry coal—				
Percentage loss in weight	9·8	8·5	7·2	8·2
	Gas 4·0, H_2O 5·8	Gas 4·3, H_2O 4·2	Gas 4·5, H_2O 2·7	Gas 4·7, H_2O 3·5
Cubic feet of gas at N.T.P. per ton	810	880	890	940
Percentage composition of the gas				
CO_2	71·6	68·3	76·9	75·7
CO	25·9	30·6	19·9	23·0
N_2	2·5	1·1	3·2	1·3
Gross calorific value of residual fuel (K.C.Us. per kgrm.)	6330	6465	6230	6130
K.C.Us. in residue from 1 kgrm. dry coal	5700	5785	5780	5490

of carbon; but, whilst carbon dioxide was still the predominating gaseous product, it was associated with a much larger proportion of the monoxide than had been found in the gases similarly obtained from any of the less mature coals previously investigated. Thus, for example, the ratio CO_2/CO in the gases from these Saskatchewan samples varied between 2·5 and 3·5, as compared with 22·0 for the Morwell coal, and between 5·0 and 12·0 for all (except one) of the Italian samples. No hydrocarbon at all was evolved during the heat treatment up to the experimental temperature limit. The carbonaceous residues obtained had the following ultimate compositions and calorific values, namely:—

Residual Fuels from Samples.

	1.	2.	3.	4.
	per cent.	per cent.	per cent.	per cent.
Carbon	68·0	68·4	67·5	67·3
Hydrogen	4·4	4·3	4·0	3·5
Nitrogen and sulphur	1·4	1·8	1·6	2·1
Oxygen	19·3	15·8	16·8	14·4
Ash	6·9	9·7	10·1	12·7
K.C.Us. per kilogramme	6380	6465	6230	6130

The results obtained with the four different samples were so nearly alike that the following composite balance-sheet for the experimental treatment may be regarded as substantially correct for the series as a whole:—

	100 parts of the original dry coal containing	} yielded {	91 parts of residual coal containing	Loss.
	parts.		parts.	parts.
Carbon	63·25		62·0	1·25
Hydrogen	4·35		3·7	0·65
Oxygen	22·20		15·6	6·60
K.C.Us.	586,000		576,000	10,000

Thus it is seen that the dry coals lost by the heat treatment about one-fiftieth part of their original carbon (as CO_2 and CO in the ratio of about 3 to 1 by volume), between one-sixth and one-seventh part of their original hydrogen (as steam), and as much as 30 per cent. of their original oxygen (the total loss in weight amounting to one-twelfth of that taken), but retained more than 98 per cent. of their potential energy in the residual fuel.

(4) *Investigation of a Burmese Lignite.*—The following details concerning the heat treatment of a Burmese lignite investigated may be cited as showing the extent to which the fuel value of a mature lignite of Class D may be improved. The sample in question had a shiny-black colour and a well-marked conchoidal fracture.

As received, it contained about 10 per cent. of moisture; the dry coal contained 4·1 per cent. of ash, and yielded at 900° C. 45·8 per cent. of volatile matter, equivalent to 47·7 per cent. on the "dry ashless" coal.

It was found that the practical limit to which the heat treatment could be pushed was 320° C., at which temperature hydrocarbon oils first appeared. Up to this limit the dried fuel suffered a loss of 7 per cent. in weight, of which 4·3 was due to gases containing 83·2 per cent. of carbon dioxide, 11·4 per cent. of carbon monoxide, and 5·4 per cent. of nitrogen.

The following was the balance-sheet of the experimental treatment:—

	100 parts of the original dry coal containing	} yielded {	93 parts of residual coal containing	Loss.
	parts.		parts.	parts.
Carbon	63·6		62·3	1·3
Hydrogen	4·8		4·1	0·7
Oxygen	26·1		20·4	5·7
K.C.Us.	612,000		604,000	8,000

C. *Summary and Conclusions.*

Having now investigated representative samples, obtained from widely different localities, of the four principal classes of tertiary brown coals and lignites, and having found no material difference between their behaviour on heat treatment within the range of temperature in question, it seems highly probable that such behaviour is common to lignites as a whole. Accordingly, it may be concluded:—

(1) that there is for each particular lignite a certain definite temperature limit (usually between 300° and 400° C.) up to which it may be heated in the dry state so as to effect a considerable chemical condensation in its cellulosic or humic constituents, with simultaneous expulsion therefrom of steam and carbon dioxide, together with a small but variable proportion of carbonic oxide;

(2) that the said chemical condensation is unaccompanied, within the temperature range in question, by any other change productive of either hydrogen or hydrocarbons;

(3) that the consequent loss in weight experienced by the "dry ashless"

coal substance up to the temperature limit in question may amount to anything between 8 and 15 per cent. of its original weight;

(4) that such loss in weight occurs principally at the expense of the oxygen-content of the dry coal, which is diminished by between one-quarter and one-third of its original value;

(5) that, substantially, the whole of the potential energy of the fuel may be correspondingly concentrated by suitable heat treatment (within the prescribed temperature limit) in the resulting carbonaceous residue, which may therefore be burnt with greater calorific intensity than the original coal; and

(6) that, accordingly, such treatment constitutes a possible means of "up-grading" brown coals and lignites generally, thus improving their fuel values.

It is not proposed, at this stage of the inquiry, to discuss the possible bearing of the experimental results on the question of coal formation, though obviously they may prove to be of considerable interest in such connection, as showing how the cellulosic or humic part of a consolidated deposit of vegetable *débris* would probably behave if subjected to the quite moderate degree of heat contemplated in the paper. At temperatures between 300° C. and 400° C., the chemical condensation in question rapidly reaches its end-point for that temperature; as a matter of fact, it would usually be completed within one or, at the most, two hours. At lower temperatures, it would, of course, proceed more slowly, but it would tend to the same end-point. Before, however, entering upon the discussion of such larger questions, it is proposed to carry out further experiments, with the object of ascertaining more precisely (*a*) the influence of pressure, as well as temperature, upon the speed and course of the chemical condensation in question, and (*b*) the cause of the variation in the proportion of carbonic oxide evolved. These and other aspects of the question are now being investigated under the author's direction at South Kensington.

In conclusion, the author desires to express his best thanks to Sir Peter MacBride (Agent-General for Victoria), the Lignite Utilisation Board of Canada, the Italian Military Mission in London, Mr. W. R. Wood and the Underfeed Stoker Company, and other friends, for their courtesy in sending him a large number of samples of lignite, from which those dealt with in this paper have been selected. He also desires to thank Messrs. C. Godfrey, A.R.C.S., and W. E. King, A.R.C.S., for their valuable assistance with the experimental part of the work.

On the Transformation of Integrals.

By W. H. Young, Sc.D., F.R.S., Professor of Mathematics in the University of Wales.

(Received January 21, 1921.)

§ 1. It is known that, if

$$x = x(u, v), \qquad y = y(u, v),$$

be one-valued continuous functions of (u, v), possessing continuous differential coefficients, while $f(x, y)$ is any continuous function of (x, y),

$$\iint_C f(x, y)\,dx\,dy = \int_a^c \int_b^d f\{x(u, v), y(u, v)\} \frac{\partial(x, y)}{\partial(u, v)}\,du\,dv, \tag{I}$$

where the integration on the left-hand side is taken over the area of the plane curve, C, which is the image in the (x, y)-plane of the rectangle $(a, b;\ c, d)$ in the (u, v)-plane. Here it is tacitly assumed that C divides the plane into two distinct parts, a limitation which, however, disappears when we employ the definition of integration over an area which I have found it necessary to introduce into analysis.*

§ 2. If we denote by $F(x, y)$ one of the indefinite integrals of $f(x, y)$ with respect to x, and by

$$x = x(t), \qquad y = y(t),$$

suitably chosen equations of the curve C, the surface integral on the left-hand side of (I) may be replaced by a contour integral (*loc. cit.*). In fact, we may then write (I) in the form

$$\int_{t_0}^{t_1} F\{x(t), y(t)\}\,dy(t) = \int_a^c \int_b^d f\{x(u, v), y(u, v)\} \frac{\partial(x, y)}{\partial(u, v)}\,du\,dv, \tag{II}$$

an identity which holds good however the curve C cuts itself, provided only the function $y(t)$ has bounded variation, or, as I have elsewhere† phrased it, provided C is a semi-rectifiable curve.

§ 3. The whole of the theory of the transformation of the variables in multiple integrals may be said to depend on identities of the forms (I) and (II). The problem of determining the most general conditions under which they hold good must be regarded as one of the most fundamental in the whole range of analysis.

---

* See a recent paper by the author "On Integration over the Area of a Curve and Transformation of the Variables in a Multiple Integral," 1920. To appear in the 'Proc. L. M. S.'

† Address to the Congrès de Mathématiciens at Strasbourg, 1920.

Employing a method explained in my paper "On a Formula for an Area,"[*] I recently obtained (*loc. cit. prim.*) such a set of conditions. Still more recently,[†] I have constructed a new method of dealing with these matters, and I propose in the present communication to obtain by means of it a second, perhaps still more strikingly, general set of sufficient conditions, involving the notion of *associated summabilities*, which I had occasion some years back to explain to the Society.[‡] This method seems of itself sufficiently remarkable to merit attention, not only for its own sake, but also for the numerous applications of which it appears to be capable.

The set of conditions (III) obtained is as follows:—

(i) *$x(u, v)$ and $y(u, v)$ are each absolutely continuous functions of each of the variables u and v separately:—*

(ii) *their partial derivates $\partial x/\partial u$, $\partial x/\partial v$, $\partial y/\partial u$, and $\partial y/\partial v$ have absolutely convergent Lebesgue double integrals, and either,*[§]—

(iiia) *in each of the pairs of functions $(\partial x/\partial u, \partial y/\partial v)$, $(\partial x/\partial v, \partial y/\partial u)$, one function is bounded, or else—*

(iiib) *the functions in each pair have associated summabilities.*

§ 4. It is of importance to remark that the theorem just stated does not cease to hold when $f(x, y)$ is no longer continuous. By employing the method of sequences, we can successively replace $f(x, y)$ by functions of increasing degree of generality, and obtain, for example, the theorem that the integral of any bounded function $f(x, y)$ over a generalised area is expressible in the form (I). For the details of the argument we refer to the paper first cited above.[||] It is not necessary to repeat it here, as the reasoning does not depend on the particular form of the conditions hypothecated for the functions $x(u, v)$ and $y(u, v)$.

In what follows, we shall accordingly suppose $f(x, y)$ continuous.

§ 6. We begin by supposing that $y(u, v)$ possesses the property of having partial differential coefficients which are continuous functions of (u, v), while $x(u, v)$ has only the properties stated in III. Writing

---

[*] 'Proc. L. M. S.,' Ser. 2, vol. 18, pp. 239–374.

[†] "A New Method in the Theory of Areas," 1920. To appear in the 'Proc. L. M. S.'

[‡] "On Classes of Summable Functions and their Fourier Series," 1912, 'Roy. Soc. Proc.,' A, vol. 87.

[§] Conditions (iii*a*) and (iii*b*) refer to each pair separately, thus, for instance (iii*a*), may hold for $\left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial v}\right)$ and (iii*b*) for $\left(\frac{dx}{\partial v}, \frac{\partial y}{\partial u}\right)$. Moreover, in different intervals these alternatives may vary, and the summabilities may also vary, provided they remain associated.

[||] See the paper on "A New Method . . ." (*loc. cit.*), where $\bar{Q}(u, v)$ is used for what is here designated $Q(u, v)$.

$$Q(u, v) = \frac{1}{hk}\int_u^{u+h}\int_v^{v+k} x(u, v)\,du\,dv = \int_0^1\int_0^1 x(u+t_1h,\ v+t_2k)\,dt_1\,dt_2,$$

and therefore

$$\frac{\partial Q}{\partial u} = \frac{1}{hk}\int_u^{u+h}\int_v^{v+k}\frac{\partial x}{\partial u}\,du\,dv = \int_0^1\int_0^1\frac{\partial}{\partial u}x(u+t_1h,\ v+t_2k)\,dt_1\,dt_2,$$

we see that $Q(u, v)$ has precisely the same properties as those we are retaining for $y(u, v)$. Thus (II) holds, if on the right-hand side we replace $x(u, v)$ by $Q(u, v)$, and at the same time make the corresponding change of $x(t)$ to $Q(t)$ on the left-hand side.

Now, if $x(u, v)$ is a continuous function of (u, v), we can determine constants h_0 and k_0, so that for *all* values of $|h| \leq h_0$ and of $|k| \leq k_0$,

$$|x(u+t_1h,\ v+t_2h)-x(u, v)| \leq \epsilon,$$

t_1 and t_2 lying between 0 and 1, whence also

$$|Q(u, v)-x(u, v)| \leq \epsilon.$$

In this case, $x(u, v)$ is the *unique double limit* of $Q(u, v)$ when $(h, k)\to(0, 0)$.

If, however, as in our conditions, $x(u, v)$ is only assumed to have the property of continuity with respect to each variable u and v separately, we obtain similarly $x(u, v)$ as the *repeated* limit of $Q(u, v)$, when either $h\to 0$, then $k\to 0$, or when first $k\to 0$, then $h\to 0$. Also, whereas in the former case, the approach to the double limit was seen to be uniform doubly, now it is uniform with respect to each variable in turn,* and we may still assume that, for all values under consideration,

$$|Q(u, v)-x(u, v)| \leq \epsilon,$$

as h and k approach their limits in the given order.

§ 7. Taking then the equations we have obtained

$$\int_{t_0}^{t_1} F\{Q(t), y(t)\}\,dy(t) = \int_a^c\int_b^d f\{Q(u, v), y(u, v)\}\frac{\partial(Q, y)}{\partial(u, v)}\,du\,dv, \qquad (1)$$

where

$$Q(t) = Q(u, v) = \int_0^1\int_0^1 x(u+t_1h,\ v+t_2k)\,dt_1\,dt_2, \qquad (2)$$

---

* For instance, taking $h\to 0$, $k\to 0$, we have for any chosen value of k a value h_{0t_2k} depending on t_2k, but independent of h or t_1, such that

$$|x(u+t_1h,\ v+t_2k)-x(u,\ v+t_2k)| < \tfrac{1}{2}\epsilon.$$

Also we can choose our value of k to be less than a certain k_0, for which

$$|x(u,\ v+t_2k)-x(u, v)| < \tfrac{1}{2}\epsilon,$$

and, therefore,

$$|x(u+t_1h_1)v+t_2k)-x(u, v)| < \epsilon,$$

and

$$|Q(u, v)-x(u, v)| < \epsilon.$$

for values of (u, v) on the perimeter of the rectangle $(a, b;\ c, d)$, and

$$|\mathrm{Q}(u, v)-x(u, v)| \leq \epsilon. \tag{3}$$

We have now to carry out and to justify the procession to the limit (repeated limit, $h \to 0$, $k \to 0$ or $k \to 0$, $h \to 0$) under the functional and integral signs, so as to obtain the equation (II).

§ 8. We commence with the left-hand side of (1).

We have, since $\mathrm{F}(x, y)$ is continuous with respect to (x, y),

$$|\mathrm{F}\{\mathrm{Q}(t), y(t)\}-\mathrm{F}\{x(t), y(t)\}| \leq e, \tag{4}$$

where e is any chosen quantity, provided, in virtue of (3), ϵ be chosen conveniently small. Therefore the left-hand side of (4), as we proceed to the limit considered, tends uniformly to zero, and, therefore, we may integrate term-by-term with respect to the function of bounded variation $y(t)$.

Hence the left-hand side of (1) tends to the required form, namely, the left-hand side of (II).

§ 9. We now proceed to discuss the right-hand side of (1). We break up the discussion into two parts, taking each term of the Jacobian separately, and proceed first to show that

$$\int_a^c \int_b^d \left\{ f(\mathrm{Q}, y) \frac{\partial \mathrm{Q}}{\partial u} \cdot \frac{\partial y}{\partial v}\, du\, dv - f(x, y) \frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v} \right\} du\, dv \to 0. \tag{5}$$

We may write the integrand here as the sum of two terms:—

$$f(x, y) \frac{\partial y}{\partial v} \left\{ \frac{\partial \mathrm{Q}}{\partial u} - \frac{\partial x}{\partial u} \right\} + \frac{\partial \mathrm{Q}}{\partial u} \cdot \frac{\partial y}{\partial v} \{ f(\mathrm{Q}, y) - f(x, y\}. \tag{α}$$

Now I have already proved in "A New Method in the Theory of Areas" that the double integral of the first of these two terms tends uniformly to zero, if the factor $f(x, y)$ be omitted. The proof, which only assumes that x and y obey the conditions (III), is not affected by the multiplication of the function $\partial y/\partial v$ by the continuous function $f\{x(u,v),\ y(,u\ v)\}$. The new function is, in fact, still independent of h and k, and has the same summability properties as $\partial y/\partial v$. Thus we may assume the proof, and write,

$$\int_a^c \int_b^d \left\{ \frac{\partial \mathrm{Q}}{\partial u} - \frac{\partial x}{\partial u} \right\} f(x, y) \frac{\partial y}{\partial v}\, du\, dv \to 0. \tag{6}$$

The second term in (α) is the product of $\partial \mathrm{Q}/\partial u \,.\, \partial y/\partial v$ into a factor $\{f(\mathrm{Q}, y)-f(x, y)\}$ which, in virtue of (3), is numerically less than e,—since $f(x, y)$ is continuous with respect to (x, y),—provided ϵ be chosen conveniently small. When integrated, therefore, the result is numerically

$$< e \int_a^c du \int_b^d dv \int_0^1 \int_0^1 \left| \frac{\partial y}{\partial v} \right| \left| \frac{\partial}{\partial u} x(u+t_1 h,\ v+t_2 k) \right| dt_1\, dt_2,$$

$$< e \int_0^1 dt_1 \int_0^1 dt_2 \int_a^c \int_b^d \left| \frac{\partial y}{\partial v} \right| \left| \frac{\partial}{\partial u} x(u+t_1 h_1,\ v+t_2 k) \right| du\, dv,$$

where we are justified in changing the order of integration, since the quadruple integral exists by reason of the association of the summabilities of $\partial y/\partial v$ and $\partial x/\partial u$.[*] Moreover, by the fundamental[†] theorem in the theory of such pairs of functions, the inside double integral is less than a finite constant B for all the values of (h, k) considered. Hence, since ϵ is as small as we please,

$$\int_a^c \int_b^d \frac{\partial Q}{\partial u} \cdot \frac{\partial y}{\partial v} \{f(Q, y) - f(x, y)\}\, du\, dv \to 0. \tag{7}$$

By (6) and (7) the double integral of (α) tends to zero, which proves (5).

Similarly the second term of the Jacobian may be treated, whence, combining the two terms, the required result follows, namely, the right-hand side of (1) tends to that of (II).

We repeat once more that the limit we have here taken is, with the conditions imposed on $x(u, v)$ in (III), the repeated limit, when first $h \to 0$, then $k \to 0$, or when first $k \to 0$, then $h \to 0$.

§ 10. We have already remarked that the theorem borrowed from "A New Method in the Theory of Areas" holds good when neither x nor y have continuous differential coefficients, but the conditions (III) hold good. We can therefore repeat the argument, and deduce the truth of (II) for the general case from that just proved, in which one of the pair (x, y) has continuous differential coefficients. All that is necessary is to make a certain change in dealing with the left-hand side of (II).

For this purpose, we remark that the left-hand side of (II) may be written

$$-\int G\{x(t), y(t)\}\, dx(t),$$

where $G(x, y)$ is an indefinite integral with respect to y of $f(x, y)$. Here the function with respect to which we integrate satisfies conditions (III), while the other function $y(u, v)$, involved implicitly in the integrand, has continuous differential coefficients. Thus, replacing $y(u, v)$ by

$$R(u, v) = \frac{1}{hk} \int_a^c \int_b^d y(u, v)\, du\, dv,$$

and carrying out precisely the same argument as in § 9, the proof of our theorem is completed.

---

[*] *Cp.* "A New Method in the Theory of Areas."

[†] *Ibid.*, proof of Lemma 3. Theorem originally given in "On Classes of Summable Functions . . .," *loc. cit.*, p. 226.

On the Moving Striations in a Neon Tube.[*]

By T. Kikuchi, Cavendish Laboratory.

(Communicated by Sir E. Rutherford, F.R.S. Received April 22, 1921.)

The following paper is the continuation of the account of experiments on the moving striations in neon tubes already communicated.[†] In the previous experiments the discharge was excited by an induction coil, but in these direct current was used and the magnitude of the current was under control.

Three Geissler tubes (A, B, C) were used, diameters of the capillary parts were respectively 2·2 mm., 1·41 mm. and 1·27 mm., their lengths being 10·8 cm., 7 cm. and 8·5 cm. These tubes (A and B) had auxiliary platinum wire electrodes at the ends of the capillary parts of the tubes to measure the drop of potential along that part.

The rotating mirror and the telescope were set practically the same as in the previous experiments. The distance from the tubes was about 2 metres and the rotation of the mirror was 200—1000 revolutions per minute. The method of measuring the velocity and the arrangement of the pump, etc., were the same as before.

A cabinet of 500 small accumulators was used as the source of the current with an adjustable water rheostat in series with the tube to control it.

Care was taken to obtain the neon gas as pure as possible, but in the gas used there was some evidence of mercury vapour and traces of hydrogen. When these impurities were less than a certain amount, no appreciable difference in the behaviour of the moving striations seemed to result from their variations. It seems probable that the effect of the impurities is rather secondary and the main results are the same as in pure neon.

The Three Types of Moving Striations.

In the case of tubes used there are three typical forms of striations all of which move toward the cathode.

The first type is what was observed in the previous experiments with an induction coil. The striations are very distinct and distant from each other, *i.e.*, in the case of tube A, about one striation per centimetre, and for tube B two or three striations per centimetre. They move with fluctuating velocity, forming a wavy pattern when observed through the rotating mirror. Also, when the current is small their mean velocity is not constant; it is less near the anode end, but when the current increases the mean velocity becomes nearly the same throughout the tube. Their velocity decreases as the pressure of the neon increases, but it is in general nearly independent of the current. In smaller tubes when the pressure is relatively low the velocity seems to decrease at first as the current increases and then becomes nearly constant. In general the velocity is larger in narrower tubes, It is usual for this type of striation to develop when the current is large, the pressure high and the diameter of the tube comparatively large. The order of velocity is 2—6×10^4 cm./sec.

---

[*] This paper was prepared by Mr. T. Kikuchi before his fatal illness.—E. R.

[†] 'Roy. Soc. Proc.,' A, vol. 98, p. 50 (1920).

The second type consists of slow moving striations with velocities of the order of 0·5—$1·5 \times 10^4$ cm./sec. They are quite distinct but not so clear as the first type. There are about three striations in tube A and four or five striations in tube B per centimetre respectively. The distance between the striations is not constant, but seems to increase with increase of the current. The velocity is uniform throughout the tube so that they look like straight parallel lines when observed through the rotating mirror. Their velocity increases as the current increases, in many cases roughly proportional to the square root of the current, but it is nearly the same for the same current in tubes A and B. This type of striation is developed when the current and the pressure are moderate.

Striations of the third type are very faint and close together, moving very fast (7 or 8×10^4 cm./sec.). The velocity of this type seems to decrease as the pressure of the gas decreases. The small current, low pressure, and narrow tube are favourable for their appearance; but as they are faint it is difficult to study them in detail. The electric force in the capillary part of the tube A was about 20 volts per centimetre, and in tube B it was about 30 volts per centimetre, and as the current was increased the electric force decreased a little, but no discontinuity was detected. The electric force was nearly independent of the pressure used in this experiment.

The Relation of the Three Types of Striations.

Take the case of fig. 1, *i.e.*, tube B, neon, 5 mm. Hg pressure. When the current is less than 1 milliampère the third type of moving striations is seen nearly throughout the tube; at about 1 milliampère the second type of striations starts from the cathode side, and as the current increases, the striations fill the whole length of the tube. At about 4 milliampères the first type of moving striations sets in from the anode side, and as the current increases, the region of the first type expands until the second type can no longer be seen in the tube.

This is the usual phenomena when three kinds of striations are seen together, but as the pressure increases (above 5 mm. Hg) the third type does not appear at all and the first type begins to appear at a smaller current than before (see fig. 2). When the pressure is still higher (above 20 mm. Hg), the second type of striations also vanish. When the pressure becomes lower, the third type seems to develop. Fig. 3 shows the appearance when the first type and the second type are observed together through the rotating mirror.

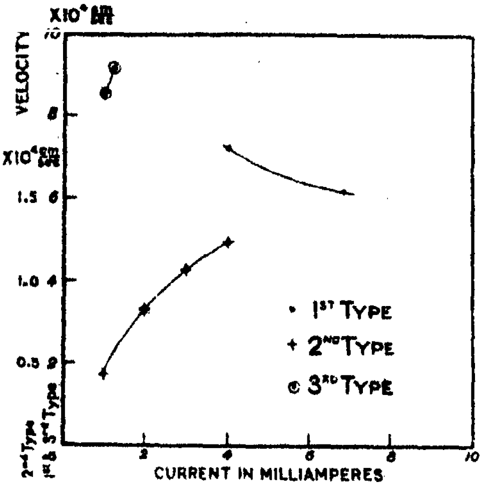


Fig. 1.

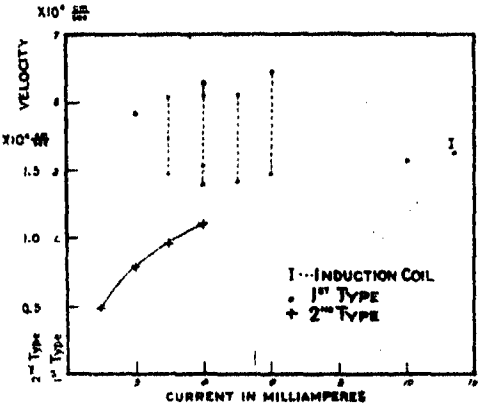


Fig. 2.

In tube A the effect of the larger bore is nearly the same as the increase of pressure in tube B. But the current at which the first type of striations sets in is nearly the same for tubes A and B, so that the current density is much larger in the narrower tube. In tube C the effect is as though the pressure is decreased in tube B, so that the third type of striations is observed still at about 100 mm. pressure. In narrower tubes, especially at low pressures, some irregularities were observed. For instance, the velocity of the second type was not very uniform throughout the tube, and showed somewhat wavy form in the rotating mirror, and their velocity increased faster than the square root of the current, and also

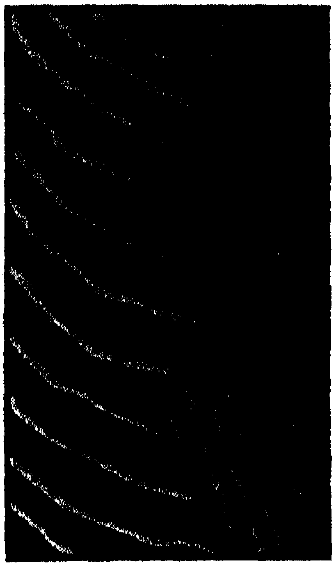


Fig. 3.

the mean velocity of the first type decreased at first as the current increased.

The Effect of Liquid Air and the Magnetic Field.

Looking at the tube with a prism, the mercury lines become fainter or disappear near the anode end, and it was thought that, as the first type of moving striations sets in from this side, the mercury vapour might be responsible for the second type of striations. Therefore, for three hours during discharge, liquid air was applied to a side tube. The mercury lines became very faint, but there was no change in the behaviour of the striations, showing that either mercury has nothing to do with the striations or a very small amount is effective.

The tube was placed between the poles of a small electro-magnet, to observe the effect of a magnetic field. When the pressure was about 5 mm. Hg, the magnetic field made the second type of striations much fainter, but the first type was not affected. Above 10 mm. Hg pressure there was no effect on either type of striation.

The Effect of Temperature.

The tube C was put inside a metallic cylinder and was heated to about 140° C. When this experiment was tried at pressures such that a slight change of pressure produced a large effect, it was clear that for the first and second types, when the pressure was kept constant, the raising of the temperature produced an effect similar to a slight decrease of pressure, and when the volume was kept constant, the effect was similar to a slight increase of pressure. The third type showed an increase in their velocity when the temperature was raised.

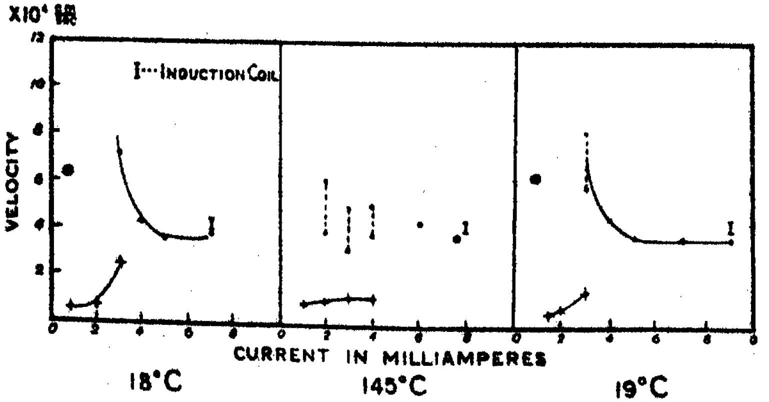


FIG. 4.

Fig. 4 shows the case where the volume was kept constant. In this case the change of the behaviour of the moving striations when the temperature was raised was conspicuous.

The author wishes to express his hearty thanks to Prof. Sir Ernest Rutherford and Dr. Aston for their kind interest and guidance.

The Magnetic Spectrum of the β-Rays Excited by γ-Rays.

By C. D. Ellis, B.A., Coutts-Trotter Student of Trinity College, Cambridge.

(Communicated by Sir Ernest Rutherford, F.R.S. Received April 22, 1921.)

In recent years, the photo-electric effect of light has been worked out in considerable detail, especially as regards the dependence of the velocity of the ejected electron on the frequency of the absorbed light, but very little knowledge exists about the analogous phenomenon in the region of higher frequencies. A very simple method of investigating the velocities of the β-rays was used by Rutherford, Robinson, and Rawlinson.* They found that the γ-rays from radium B and C, when passed through heavy metals, such as gold or lead, caused the emission of several groups of electrons, each group consisting initially of electrons of the same velocity. They showed that this fact alone gave information about the connection between the β-rays and γ-rays of radio-active bodies, but, in addition, they made the interesting observation that the velocity of the electrons liberated from gold had 1 to 2 per cent. higher velocity than those from lead.

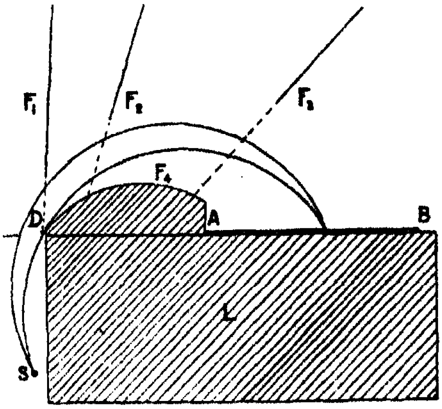
This fact receives a simple explanation if it be assumed that the energy of the emitted electron is equal to some energy characteristic only of the γ-ray, minus the energy necessary to remove the electron from the atom. The difference in the energies of the electrons ejected from gold and lead by the same γ-rays is then explained by the difference in the work of removal of these electrons from their respective atoms. Since the general relation of the gold atom to the lead atom is known, it should be possible to deduce from the experimental values for this difference in the work of removal, from what part of the atom the electron originated, and then values might be obtained for the energy characteristic of the γ-ray. The work to be described was undertaken with these points in view, the experimental determination consisting in the measurement of the energies of the electrons ejected from various metals by the γ-rays of radium B and C, with the special object of measuring accurately the differencies of these energies from metal to metal. As wide a range of metals as was practicable was used, in order to test any explanation of these energy differences.

---

* 'Phil. Mag.' (2), p. 277 (1914).

The Experimental Method.

The experimental method used was identical with that of Rutherford, Robinson, and Rawlinson (*loc. cit.*), in which the energies of the electrons are measured by bending them in a magnetic field. By a method to be described later, a line source of the electrons to be investigated was obtained. This source, S (see fig.), was fixed about 3 mm. from the end of a lead block, L (10·5 × 4·5 × 3 cm.). A comparatively wide slit, D (4–7 mm.), was arranged vertically above the source, and the latter was usually about 3·6 cm. below the slit. A photographic plate, AB, is fixed in position on the top of the lead block. The whole apparatus is lowered into a brass box and exhausted to a low vacuum. The box is then placed between the poles of a large electromagnet, which are so adjusted as to give a nearly uniform field over the space between them, the magnetic field being parallel to the source and the slit. In the uniform magnetic field, electrons of a definite velocity describe circles of the same radius, and, even with the wide slits used, the circles intersect the photographic plate along a curved line of comparatively narrow width. The existence of a group of electrons of definite velocity is thus shown by a line on the plate.



The general appearance of the photographs was the same as described by Rutherford (*loc. cit.*); the outside edges of the lines were sharp, and the radius, ρ, of the circular path of the rays forming this part of the trace was determined by the formula $\rho = \sqrt{(D^2+a^2)}$, where $2a$ is the distance of the source from the slit and 2D is the distance of the outside of the line from the vertical through the top of the source. The dimensions of the apparatus are larger than that used by Rutherford, ρ varying between 3 and 4·5 cm., since, although this diminishes the absolute intensity of any line, it increases the contrast against the general background. Also, slightly more elaborate screening arrangements were adopted, screens $F_1F_2F_3F_4$ of lead being used to cut down stray radiation.

The source of γ-rays, S (see fig.), was a small tube about 10 mm. long and 0·7 mm. diameter, filled with radium emanation. A thin sheet of the metal to be investigated was wrapped tightly round the tube. Radium B and C are deposited from the emanation on the walls of the tube, and the γ-rays from these passing out of the tube traverse the metal and liberate high-speed electrons from it. In this way an approximately line source is obtained, and the nearness of the radiator to the radiation gives the maximum intensity. Five elements were investigated, Ur, Pb, Pt, W, and Ba. Pb, and Pt were used in the form of foils, 0·3 mm. thick, as described; the other three were taken in the form of oxides. In this case the emanation tube was covered with a silver sheath, and a thin layer of the oxide was fixed on to the silver. Special experiments showed no intense lines were obtained from the silver.

The values of H, the magnetic field, were determined by taking photographs with the same apparatus of the natural β-ray spectrum of radium B, and using the values of $H\rho$ for these lines found by Rutherford and Robinson.* In place of the usual active wire, a plate source was used (1 × 0·3 cm.). If this is tilted so that a line perpendicular to its plane passes somewhere in the region of A (see the fig.), excellent focussing is obtained. Such a source is far more convenient to prepare than an active wire.

$H\rho$ is determined experimentally, and the energy corresponding to this is obtained by means of the relativity formulæ. It is found very convenient to characterise these energies by stating the potential difference in volts an electron would have to fall through to acquire this energy.

The region investigated is that corresponding to the main β-ray emission of radium B, that is, to β-rays of speeds 0·8 to 0·6 of the velocity of light.

Experimental Results.

With each of the elements, tungsten, platinum, lead, and uranium, a group of three main lines was found. Other fainter lines were present, but these strong lines will be considered first. Each group had the same general appearance, the only difference being that each of the three lines from a metal of smaller atomic number were shifted by the same amount in the direction of increasing energy in relation to the corresponding lines from a metal of higher atomic number. This can be seen from Table I, in which under each element are arranged the energies of the three lines comprising the group: corresponding lines for different elements are in a horizontal row.

---

* 'Phil. Mag.,' vol. 26, p. 717 (1913).

Table I.

Metal	Barium.	Tungsten.	Platinum.	Lead.	Uranium.
Atomic No.	56	74	78	82	92
Energies in volts × 10^{-5}	--	1·66	1·58	1·49	1·22
	2·53	2·20	2·12	2·03	1·74
	--	2·76	2·69	2·60	2·31

Thus the tungsten lines are all about 8000 volts higher energy than the platinum lines. Only one line was measured for barium, so it cannot be used in this connection.

It is clear from this that there are three distinct γ-rays acting, each being responsible for one line, and further that the energies observed can all be included in a statement as follows:—

If E = energy of the ejected electron,

W_1, W_2, W_3, represent energies characteristic of the three γ-rays and independent of the atom acted upon, and

W_a represent an energy characteristic of the atom,

then for any one element

$$E = W_n - W_a, \qquad n = 1, 2, 3.$$

For another element W_a will have a different value, and a fresh group of lines will be obtained.

The obvious interpretation of W_a is that it represents the energy necessary to remove the electron in question from its position inside the atom to the surface. From the magnitude of the change of W_a from atom to atom (8000 volts between platinum and tungsten) it will be seen that the electrons giving these lines must come from the K rings. The energy necessary to remove an electron from the K ring is obtained from the wave-length λ of the K absorption band by the quantum relation $E = hc/\lambda$, and is expressed in volts V by the relation $V = E/e$, where e is the electronic charge.

The experimental values of the wave-lengths of the K absorption bands and the corresponding energy in volts are shown in Table II with the

Table II.

Metal.	λ in Å.U.	V in volts.	References.
Barium	0·3306	$0 \cdot 378 \times 10^5$	Siegbahn and Jönsson.[*]
Tungsten	0·1781	0·693	Duane and Patterson.[†]
Platinum	0·1578	0·782	Sieg and J.
Lead	0·1885	0·891	Sieg and J.
Uranium	0·1048	1·178	Sieg and J.

[*] ‘Phys. Zeit.,’ vol. 20, p. 251 (1919).

[†] ‘Phys. Rev.,’ p. 526 (December, 1920).

To each of the energies of the ejected electrons (see Table I) is now added the energy necessary to remove an electron from the K ring and the values obtained are shown in Table III.

Table III.

Barium.	Tungsten.	Platinum.	Lead.	Uranium.
—	2·35	2·36	2·38	2·40
2·90	2·89	2·91	2·92	2·92
—	3·46	3·46	3·49	3·48

The numbers in a vertical row represent respectively W_1, W_2, W_3, the energies characteristic of the γ-rays, and so the numbers in a horizontal row should be constant. This is seen to be the case with fair accuracy, but the variations are mostly in one direction. It is difficult to be sure of this as the variation is within the limit of error of measurement, and in addition there is a strong possibility of a systematic error due to the tungsten lines all being weaker and less sharply defined. In any case, unless the whole phenomenon is radically different from that pictured, this steady variation, if it exists, would be due to a slight dependence of W_1, W_2, W_3 on the atom in which they were absorbed. These results show definitely that the main β-ray emission under these conditions is due to three main γ-rays, and that the β-rays emitted come from the K rings.

Before treating the fainter lines it is convenient to consider the natural β-ray spectrum of radium B. Rutherford, Robinson and Rawlinson (*loc. cit.*) showed that the magnetic spectrum of the β-rays excited in lead was identical with the natural β-ray spectrum of radium B. This result was repeated in the present work. Since radium B and lead are isotopic, and as has just been shown, the emission is due to the ejection of electrons from the electronic system of the atom, this is in complete agreement with Rutherford's theory of the connection between the β-ray and γ-ray emission of radio-active atoms.

According to his view the primary phenomenon is the emission of a β-particle from the nucleus. It may happen, through some mechanism at present unknown, that this β-particle gives rise to a γ-ray. This γ-ray in traversing the electronic system of the atom may be absorbed and eject a high speed electron. It is these last electrons which constitute the β-ray line spectrum.

As a result of this present work it is possible to amplify this theory in some details. The evidence just brought shows that the three main β-ray lines originate in the K ring due to the action of three γ-rays, and proceeding in this way it can be shown that all of the faster lines of the β-ray spectrum of radium B, both strong and faint, can be accounted for by supposing that there are six γ-rays, each characterised by a definite amount of energy, and that each of these γ-rays ejects electrons from one of two definite levels in the atom. The first of these is the K ring and the second the Λ' level* which is one of the L group. To explain the finer details of the L lines in the X-ray spectra of the elements it is found necessary to assume the existence of four distinct levels. In the order of increasing energy they are called $LL'\Lambda\Lambda'$. Absorption bands corresponding to these levels have been obtained and their existence seems assured.

For atoms of atomic number 82 (lead, radium B) the wave-lengths of the L and L′ absorption bands were found by De Broglie† to be 0·945 Å.U. and 0·811 Å.U., and the corresponding energies to remove an electron from these levels are $0{\cdot}131 \times 10^5$ volts and $0{\cdot}152 \times 10^5$ volts. For the present purpose the L′ and Λ levels can be considered identical and the difference of energy between the Λ and Λ' level is found from the magnitude of this Λ doublet in the line spectrum (Sommerfeld‡). The energy of the level then comes to $0{\cdot}157 \times 10^5$ volts. The calculation of the radium B β-ray spectrum now proceeds as follows. The values of the γ-ray energies that are assumed are shown in column I of Table IV, column II gives the wave-length in Ångstrom units deduced from the quantum relation. The justification of this is given later.

From each of these γ-ray energies is subtracted in turn the K absorption energy ($0{\cdot}891 \times 10^5$ volts) and the Λ' absorption energy ($0{\cdot}157 \times 10^5$ volts), and the results are shown in Table V.

---

* Sommerfeld, 'Atombau und Spektrallinien,' p. 180.

† Sommerfeld, *loc. cit.*, p. 206.

‡ Sommerfeld, *loc. cit.*, p. 186.

Table IV.—γ-rays of RaB.

I.	II.
Energies in volts.	λ in A.U.
$4{\cdot}000 \times 10^5$	0·0308
3·639	0·0339
3·492	0·0354
2·918	0·0423
2·529	0·0488
2·385	0·0519

Table V.—β-ray Spectrum RaB.

I.	II.	III.	IV.
Origin of Line.	Intensity.	Energy observed.	Energy calculated.
		volts.	
Λ'	*f.*	$3{\cdot}851 \times 10^5$	$3{\cdot}843 \times 10^5$
Λ'	*m.s.*	3·473	3·482
Λ'	*s.*	3·328	3·335
K	*v.f.*	3·103	3·109
Λ' K	*m.s.*	2·756	2·761 2·748
K	*v.s.*	2·608	2·601
Λ'	*m.*	2·364	2·372
Λ'	*m.s.*	2·227	2·228
K	*v.s.*	2·031	2·027
K	*m.s.*	1·645	1·638
K	*v.s.*	1·494	1·494

Column I shows the origin of the line, column II the relative intensity, column III the observed energy, column IV the calculated energy. The observed energies are slightly different from the values given in the original paper, since newer values of the constants have been used. They were calculated afresh from the experimentally determined Hρ's.

The agreement is seen to be well within the experimental error (1/300), also all the lines are given without any gaps and no extra ones are predicted. Of course, three of the γ-rays are the same that gave the groups of lines first discussed, but the values obtained from the radium B lines will be used, since these measurements are more accurate.

This explanation of the radium B β-ray spectrum might seem forced in one particular, in that half of the lines are supposed to originate in the Λ' level, a level that from X-ray work seems one of the less important of the main L group. But evidence can be brought to show that under certain circumstances the Λ′, L′, and L levels can all be detected, and that the lines originating in the Λ′ level are the strongest. This evidence is obtained from a study of the low velocity β-rays from thorium D. These lines have been measured by Hahn,* who found a strong line of velocity 0·29 and a weaker line of velocity 0·38. It is possible to obtain much greater detail with the focussing arrangement, and seven lines have been measured by the writer in this region. There are three main groups, each group consisting of a strong line with companions. The values obtained are shown in Table VI; column 1 gives the value of Hρ, column 2 the intensity, column 3 the corresponding energies, column 4 the level from which the β's originate, and column 5 the calculated energies. Three γ-rays of characteristic energies $0{\cdot}412 \times 10^5$ volts, $0{\cdot}532 \times 10^5$ volts, and $0{\cdot}564 \times 10^5$ volts are postulated, and the values for the energies of the Λ′, L′, and L ring are taken as those for a body of atomic number 82, that is, respectively, $0{\cdot}157 \times 10^5$, $0{\cdot}152 \times 10^5$, and $0{\cdot}131 \times 10^5$ volts.†

Table VI.—β-rays ThD.

1.	2.	3.	4.	5.
Hρ.	Intensity.	Energy observed.	Origin of line.	Energy calculated.
		volts.		
543	*v.s.*	$0{\cdot}254 \times 10^5$	Λ′	$0{\cdot}255 \times 10^5$
550	*s.*	0·261	L′	0·260
570	*m.*	0·280	L	0·281
663	*v.s.*	0·375	Λ′	0·375
674	*m.*	0·400	L	0·401
690	*s.*	0·406	Λ′	0·407
696	*m.s.*	0·412	L′	0·412

The agreement is seen to be good, but there are several points to notice. Firstly, the lines from the Λ′, L′, and L levels are in the order of decreasing intensity, thus the third line from the hardest γ-ray was not found, probably because it was too faint. Secondly, these γ-rays are too soft to eject electrons from the K ring, but all the levels of the L group have been found (the Λ level would be quite indistinguishable from the L′ level). The reason why the L′ line was not found in the 663 Hρ group is because the extreme intensity of the 663 line would mask a close companion line.

These measurements are given here to show, from the point of view of the ejection of electrons by γ-rays, that the Λ' level is the most important one of the L group, and, therefore, the position given to this level in the explanation of the natural β-ray spectrum of radium B is justified.

---

* 'Phys. Zeit.,' vol. 12, p. 273 (1911).

† It is not known whether these β-rays are emitted by ThC(83) or ThD(81). In either case there is no doubt that the electrons originate from the levels given.

From the results given, it is seen that the magnetic spectrum of the high-speed electrons excited by γ-rays in metals like platinum or uranium is exactly similar to the natural β-ray spectrum of radio-active bodies, the only difference being that the same radio-active atom which produces the γ-ray also probably absorbs it. Further, the main lines of the phenomenon are sufficiently described by combining with the known electronic structure of the various atoms, the assumption of certain γ-rays characterised by definite amounts of energy.

The faint lines, previously mentioned, that were observed from W, Pt, Pb, and Ur, all fall under this scheme, but it is difficult to make the measurements accurately, owing to the faintness of the lines, and it is not considered worth while giving the values.

Discussion of Results.

Evidence has been brought to show the existence of several γ-rays, each γ-ray characterised by a certain energy which is set free when it is absorbed. On the basis of the quantum theory, one would expect this energy to be hν, where ν is the frequency of the radiation. In this way the wave-lengths of these γ-rays can be calculated, and the values obtained are shown in column 2 of Table IV. These wave-lengths are shorter than any measured by Rutherford and Andrade, but γ-rays of higher frequency are known to be present. Rutherford,[*] from considerations of the absorption coefficients of γ-rays compared to absorption coefficients of X-rays generated at high voltages, showed that the wave-lengths of the main γ-rays were much shorter than had previously been supposed, and that the hardest γ-rays of radium C might be of the order 0·007 Å.U. His general results also pointed to the fact that the groups of β-rays are due to the transformation of the energy of the γ-rays in single quanta, according to the relation E = hν.

It is seen therefore, that there is good evidence to believe that the wave-lengths are right, but it should be pointed out that small corrections may have to be applied in the light of later work. In this deduction it is assumed that the law

$$\text{Energy of the electron} = h\nu - w_0$$

is exact. This has been proved to hold with considerable accuracy for light, but in the true equation there may be small correction terms which, while quite negligible for low frequencies, may have an effect when the frequency is ten thousand times greater. It may be that the phenomenon is not correctly described by an energy equation at all, and owing to the small variation of atomic number that was practicable, these experiments were unable to test this point. It is always possible, owing to the simpler relations that all dynamical quantities bear to each other at low speeds, compared to the same relations at high speed, that the energy expression might be quite accurate at low frequencies but in error at high ones. These points could be tested if the wave-lengths of the γ-rays from thorium D could be measured, and, although they are short, this should be possible by Rutherford's transmission method. If the values calculated from the characteristic energies obtained from the β-ray spectrum agreed with the values found directly it would show that the corrections considered above are not serious.

---

[*] 'Phil. Mag.,' vol. 34, p. 153 (September, 1917).

Owing to the meagre material it is not thought advisable to discuss the numerical values at this stage, but it is interesting to note that part at least of the γ-radiation is monochromatic. It has been found possible to connect these γ-rays of radium B with the γ-rays measured by Rutherford and Andrade by the crystal method, and by means of this fact it can be shown that it is probable that the quantum dynamics holds for that part of the atom from which these γ-rays come. This evidence will be given in detail later.

The different intensities of the lines present points of great interest, for in all cases it is the most firmly bound electron which gives the most intense lines. The hard γ-rays of radium B are much more absorbed by the K electrons than by the L group, and the peripheral electrons have a comparatively negligible effect. This is seen very clearly in the thorium D lines, the strongest lines come from the Λ' level, the next strongest from the L′ level, and the weakest from the L level. Even these much softer γ-rays appear to have little effect on the outer electrons.

The results as a whole suggest that the chance of an electron absorbing γ-radiation depends on the relation between the energy of the binding of the electron and the energy characteristic of the γ-ray, and that the probability of absorption becomes very small when the energy of the binding is small compared to the γ-ray energy.

It is intended to continue this work by investigating the magnetic spectrum excited in metals by the hard γ-rays of radium C and the softer γ-rays of radium B.

When this work was completed my attention was directed to some recent publications of De Broglie* on the velocities of electrons set free from metals by X-rays. Quantitative results have not yet been given, but apparently the

---

[*] 'Comptes Rendus,' Nos. 5, 9, and 12, vol. 172 (1921).

effects observed are of the same general type that have been described in this paper.

Summary.

1. The magnetic spectrum of the β-rays excited by the γ-rays of radium B in Ur, Pb, Pt, W, and Ba, has been measured.

2. It is shown that the main lines are formed by electrons ejected from the K ring by definite γ-rays, each γ-rays being characterised by a certain energy.

3. The magnetic spectrum of the β-rays of radium B is accounted for on this basis.

4. The magnetic spectrum of the low velocity of β-rays of thorium D has been measured and is explained in the same way as the radium B β-ray spectrum.

5. On the basis of the quantum theory the wave-lengths of the γ-rays of radium B are calculated from the characteristic energies found.

6. Evidence is given to show that the greater the energy of binding of an electron the greater is the probability of it absorbing this hard γ-radiation.

In conclusion I would like to express my indebtness to Prof. Sir Ernest Rutherford, who suggested the problem and the method of attack, and directed the whole course of the work. My thanks are due to Mr. G. R. Crowe for the preparation of the active sources used.

The Vibrations of a Spinning Disk.

By H. Lamb, F.R.S., and R. V. Southwell.*

(Received April 6, 1921.)

1. This paper treats of the transverse vibrations of a circular disk of uniform thickness rotating about its axis with constant angular velocity. The problem has a practical bearing, as throwing light on the occasional failure of turbine disks. In an impulse turbine blades are fitted to the rim of a thin disk, and steam is admitted to them from nozzles which are usually arranged symmetrically, but not always continuously, round the periphery. The failures appear to be sometimes due to the blades coming in contact with the adjacent parts. This points to transverse vibrations, and there can be little doubt that the phenomenon is one of resonance between the periodic forces exerted by the steam jets and the periods of free vibration of the disk.†

For a thorough investigation it would be necessary to ascertain, to a reasonable degree of approximation, the free periods of a symmetrical disk of any given profile. This may form the theme of a subsequent investigation by one of the authors. In the meantime the study of the simpler case now discussed illustrates the physical aspects of the problem, and has suggested an important simplification in method.

2. The special difficulty of the problem arises from the fact that centrifugal force and flexural rigidity are both operative, and that the normal modes of vibration depend on their relative importance. The vibrations of a non-rotating disk were fully discussed by Kirchhoff in two well-known papers.‡ In the other extreme, where the disk is so thin and the rotation so rapid that the flexural forces are negligible, the problem admits of a very simple solution, as follows:—

When an elastic disk rotates in its own plane, in relative equilibrium, the principal tensions are along and perpendicular to the radius vector. If a be the radius, ρ the density, and ω the angular velocity of rotation, their values at a distance r from the centre are*

---

* My attention was drawn to this question by Prof. Gerald Stoney, who showed me some pretty experiments on a rotating india-rubber membrane. This led to various calculations. I found subsequently that Mr. Southwell had become interested in the same problem, and a comparison showed that the methods and the results were closely parallel. This paper is accordingly a joint production, but it should be stated that the important remark that the formula $p^2 = p_1^2 + p_2^2$ of §4 gives not merely an approximation but a *lower limit* to the frequency is due to Mr. Southwell.—H. L.

† A paper by K. Baumann, read before the Institution of Electrical Engineers on March 15, 1921, gives examples of failures which have been traced to this source.

‡ "Ueber die Schwingungen einer elastischen Scheibe," 'Crelle's Journal,' vol. 40 (1850), and 'Pogg. Ann.,' vol. 81 (1850).

$$P = A(a^2-r^2)\rho\omega^2, \qquad Q = (Aa^2-Br^2)\rho\omega^2, \tag{1}$$

respectively, where

$$A = \tfrac{1}{8}(3+\sigma), \qquad B = \tfrac{1}{8}(1+3\sigma), \tag{2}$$

σ denoting Poisson's ratio. These values are obviously unaffected by a small displacement (w) normal to the plane of the disk.

The equation of transverse motion of an element of area whose polar co-ordinates are (r, θ) is, under the condition stated,

$$\rho r\delta\theta\delta r\frac{\partial^2 w}{\partial t^2} = \frac{\partial}{\partial r}\left(\mathrm{P}r\delta\theta\frac{\partial w}{\partial r}\right)\delta r + \frac{\partial}{\partial\theta}\left(\mathrm{Q}\delta r\frac{\partial w}{r\partial\theta}\right)\delta\theta, \tag{3}$$

or

$$\rho\frac{\partial^2 w}{\partial t^2} = \frac{1}{r}\frac{\partial}{\partial r}\left(\mathrm{P}r\frac{\partial w}{\partial r}\right) + \frac{\mathrm{Q}}{r^2}\frac{\partial^2 w}{\partial\theta^2}. \tag{4}$$

Assuming that w varies as $\cos s\theta\cos p_1 t$, where s is integral, we find

$$\mathrm{A}(a^2-r^2)\left(\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}\right) - 2\mathrm{A}r\frac{\partial w}{\partial r} + \left(\frac{p_1^2}{\omega^2}+s^2\mathrm{B}-\frac{s^2\mathrm{A}a^2}{r^2}\right)w = 0. \tag{5}$$

This is readily integrated by a series. Assuming

$$w = \Sigma\mathrm{C}_k\left(\frac{r}{a}\right)^k, \tag{6}$$

we have

$$\frac{\mathrm{C}_{k+2}}{\mathrm{C}_k} = \frac{k(k+2)\mathrm{A}-s^2\mathrm{B}-p_1^2/\omega^2}{\{(k+2)^2-s^2\}\mathrm{A}}, \tag{7}$$

or, if we write

$$\frac{p_1^2}{\omega^2}+s^2\mathrm{B} = m(m+2)\mathrm{A}, \tag{8}$$

$$\frac{\mathrm{C}_{k+2}}{\mathrm{C}_k} = \frac{(k-m)(k+m+2)}{(k+2-s)(k+2+s)}. \tag{9}$$

The solution which is finite for $r = 0$ begins with $(r/a)^s$. Hence, inserting the factors temporarily omitted,

$$w = \mathrm{C}\left(\frac{r}{a}\right)^s\left\{1+\frac{(s-m)(s+m+2)}{2(2s+2)}\frac{r^2}{a^2} + \frac{(s-m)(s-m+2)(s+m+2)(s+m+4)}{2.4(2s+2)(2s+4)}\frac{r^4}{a^4}+\ldots\right\}\cos s\theta\cos p_1 t. \tag{10}$$

The series in { } is hypergeometric. In Gauss's notation,

$$w = \mathrm{C}\left(\frac{r}{a}\right)^s\mathrm{F}\left(\alpha, \beta, \gamma, \frac{r^2}{a^2}\right)\cos s\theta\cos p_1 t, \tag{11}$$

where

$$\alpha = \tfrac{1}{2}(s-m), \qquad \beta = \tfrac{1}{2}(s+m+2), \qquad \gamma = s+1. \tag{12}$$

---

* Love, 'Elasticity,' §102(*b*).

The expansion is obviously convergent for $r < a$, but since

$$\gamma = \alpha + \beta, \tag{13}$$

it becomes, by a known theorem, logarithmically infinite at the edge $r = a$, unless it terminates.[*] Hence we must have

$$m = s + 2n, \tag{14}$$

where n is a positive integer. Thus

$$w = C\left(\frac{r}{a}\right)^s \left\{1 - \frac{n(n+s+1)}{1(s+1)}\frac{r^2}{a^2} + \frac{(n-1)n(n+s+1)(n+s+2)}{1.2(s+1)(s+2)}\frac{r^4}{a^4} - \ldots\right\} \cos s\theta \cos p_1 t. \tag{15}$$

Arbitrary constants may of course be added to θ and t.

The mode of vibration thus found is characterised by n nodal circles and s equidistant nodal diameters. The corresponding frequency is determined by

$$\frac{p_1^2}{\omega^2} = (s+2n)(s+2n+2)A - s^2 B, \tag{16}$$

where A and B have the values (2). That this ratio should be independent of the diameter and thickness of the disk might have been anticipated from the theory of dimensions.

The modes for which $s = 0$ involve displacement of the centre of the disk,[†] and are therefore not relevant to the case of a disk mounted on a shaft. The latter remark applies also to the mode where $s = 1, n = 0$; the disk then remains plane in form, but is slightly untrue to the axis. The most interesting modes, and the ones most easily realised experimentally, are those in which there are no nodal circles. Putting $n = 0$, we have

$$w = C\left(\frac{r}{a}\right)^s \cos s\theta \cos p_1 t, \tag{17}$$

with

$$\frac{p_1^2}{\omega^2} = \tfrac{1}{4}(1-\sigma)s^2 + \tfrac{1}{4}(3+\sigma)s. \tag{18}$$

When $s = 2$, we have two nodal diameters at right angles, and

$$\frac{p_1^2}{\omega^2} = \tfrac{1}{2}(5-\sigma). \tag{19}$$

For an india-rubber membrane we may put $\sigma = \frac{1}{2}$, practically, whence $p/\omega = 1{\cdot}5$.

---

[*] *Cf.* Rayleigh, 'Theory of Sound,' §336.

[†] It may be noticed that if we put $r/a = \sin\frac{1}{2}\phi$ the formula (15) may in this case be written

$$w = CP_n(\cos\phi)\cos p_1 t,$$

where P_n is the zonal harmonic of order n.

Writing (16) in the form

$$p_1^2 = \lambda\omega^2, \tag{20}$$

the values of λ for various modes are given in the following table, on the hypothesis $\sigma = 0{\cdot}3$, which corresponds fairly to the case of a steel disk. The mode for which $n = 0$, $s = 2$, is the most important from the practical point of view, since it is the gravest type of free vibration if we exclude (for the reason already given) the case of $n = 0$, $s = 1$.

n.	$s = 1$.	$s = 2$.	$s = 3$.
0	1	2·35	4·05
1	5·95	8·95	12·3
2	14·2	18·85	23·85
3	25·75	32·05	38·7

It is to be observed, throughout this paper, that the vibrations considered are vibrations relative to the undisturbed disk; in particular, the nodal lines in the modes investigated are in reality lines of rest in the disk, and are accordingly carried round by the rotation. The results may, however, be stated in a different form. By superposition of two normal modes of the same type, with a suitable difference of phase, we can construct wave-forms which travel round the disk with the angular velocity $\pm p_1/s$. The preceding results would, in fact, have been unaltered if we had assumed that w varies as $\cos(p_1t \mp s\theta)$ instead of $\cos p_1t \cos s\theta$. The corresponding angular wave-velocities *in space* are $\omega \pm p_1/s$. Thus, in the case $n = 0$, $s = 1$, where $p_1 = \omega$, by (16), the wave-velocities are 2ω and 0. In the latter alternative, the membrane merely rotates in a fixed plane slightly inclined to that which we have taken as the plane of reference.

3. The preceding may be regarded as results which are approximated to as the speed of rotation is increased. In the other extreme, where the influence of the rotation is negligible, and the flexural rigidity predominant, we may have recourse to the investigation of Kirchhoff. The results are of the form

$$p_2^2 = \mu \frac{\mathrm{E}h^2}{\rho a^4}, \tag{21}$$

where E is Young's modulus, h is the half-thickness, and μ is a numerical coefficient depending on the particular mode of vibration and (to some extent) on the value of σ. The results are actually presented by Kirchhoff somewhat differently. He gives tables of frequencies of various modes relative to that of the gravest mode of all as a standard, first on Poisson's hypothesis of $\sigma = \frac{1}{4}$, and secondly for $\sigma = \frac{1}{3}$. The gravest mode in question has two nodal diameters and no nodal circle. The frequency for this is given by Kirchhoff in semi-vibrations per second, viz.,

$$\frac{p_2}{\pi} = \sqrt{\left(\frac{\mathrm{E}h^2}{\rho a^4}\right)} \times 1{\cdot}04604, \tag{22}$$

and

$$\frac{p_2}{\pi} = \sqrt{\left(\frac{\mathrm{E}h^2}{\rho a^4}\right)} \times 1{\cdot}02357, \tag{23}$$

for $\sigma = \frac{1}{4}$ and $\sigma = \frac{1}{3}$, respectively.

An independent calculation for the case of $\sigma = 0{\cdot}3$ gives the following values of the coefficient μ in (21). The number of nodal diameters is denoted, as before, by s, and that of nodal circles by n.*

n.	$s = 0$.	$s = 1$.	$s = 2$.	$s = 3$.
0	—	—	10·61	56·87
1	29·67	154·3	443·9	1040
2	577	1222	2624	4285
3	2983	5266	8574	13287

4. When the centrifugal tensions and the flexural stiffness are both effective, the difficulties of an exact solution are much increased. The equation of motion is

$$\rho\frac{\partial^2 w}{\partial t^2} = \frac{1}{r}\frac{\partial}{\partial r}\left(\mathrm{P}r\frac{\partial w}{\partial r}\right) + \frac{\mathrm{Q}}{r^2}\frac{\partial^2 w}{\partial \theta^2} - \frac{\mathrm{E}h^2}{3(1-\sigma^2)}\nabla^4 w; \tag{24}$$

and the boundary conditions are as given in Rayleigh's "Theory of Sound," § 215. If we assume that w varies as $\cos s\theta \cos pt$, as before, the equation can be integrated by a series, but the relation between successive coefficients consists of *three* terms, and the deduction of general results would not appear to be easy.

Approximations which are adequate from the practical point of view can, however, be obtained by Rayleigh's method of an assumed type,† at all events, for the more important modes. The procedure is equivalent to this, that we calculate the potential and kinetic energies on the basis of an assumed form of w, and determine the frequency so that the mean energies of the two kinds shall be equal. Since the frequency thus determined is stationary for slight variations of the type, a good approximation can be found in this way if the type be suitably chosen. It is known, moreover, that in the case of the gravest mode of all, the frequency obtained will be an upper limit to the true value. In the present problem, this statement obviously applies also to the gravest mode having any assigned number s of nodal diameters.

---

* The authors are indebted to Miss B. S. Gough for assistance in the rather laborious calculations.

† 'Theory of Sound,' §§88, 89.

The potential energy of the centrifugal forces is

$$V_1 = h \iint \left\{ P \left(\frac{\partial w}{\partial r}\right)^2 + Q \left(\frac{\partial w}{r\partial \theta}\right)^2 \right\} r\, d\theta\, dr, \tag{25}$$

where P and Q have the values (1), and the integration extends over the area of the disk.

The potential energy of flexure is, in Cartesian co-ordinates,*

$$V_2 = \frac{Eh^3}{3(1-\sigma^2)} \iint \{ (\nabla^2 w)^2 + 2(1-\sigma)(w_{xy}{}^2 - w_{xx} w_{yy}) \}\, dx\, dy. \tag{26}$$

To transform to polar co-ordinates we have

$$\nabla^2 w = \frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{1}{r^2}\frac{\partial^2 w}{\partial \theta^2}, \tag{27}$$

$$w_{xy}{}^2 - w_{xx} w_{yy} = \left(\frac{1}{r}\frac{\partial^2 w}{\partial r \partial \theta} - \frac{1}{r^2}\frac{\partial w}{\partial \theta}\right)^2 - \frac{\partial^2 w}{\partial r^2}\left(\frac{1}{r^2}\frac{\partial^2 w}{\partial \theta^2} + \frac{1}{r}\frac{\partial w}{\partial r}\right). \tag{28}$$

The kinetic energy is

$$T = \rho h \iint \left(\frac{\partial w}{\partial t}\right)^2 r d\theta\, dr. \tag{29}$$

If, then, we assume

$$w = \phi(r, \theta) \cos pt, \tag{30}$$

the method referred to leads to a result which may be written

$$p^2 = \frac{V_1(\phi) + V_2(\phi)}{T(\phi)}, \tag{31}$$

where $T(\phi)$ differs from T in that $\partial w/\partial t$ is replaced by ϕ.

In estimating in this way the frequency of the gravest mode for a given value of s, it is sometimes convenient to adopt for ϕ a form containing a variable parameter, and to adjust this afterwards so as to make the fraction in (31) a minimum. This reduces the upper limit referred to. This plan is adopted, and illustrated by numerical examples, in § 5 below. The work is somewhat intricate, but the formula (31) suggests a simpler process, which is probably adequate for practical purposes, having regard to the fact that the elastic-constants E and σ for a given disk are not likely to be known to a high degree of accuracy.

It is evident that if precisely the right form be given to ϕ the formula (31) will be exact. Now, on account of the stationary property referred to, the true value of ϕ would give an approximate value for $p_1{}^2$, viz.,

$$p_1{}^2 = \frac{V_1(\phi)}{T(\phi)}, \tag{32}$$

---

* 'Theory of Sound,' §214.

In the same way the true value of ϕ would give an approximate value for p_2^2, viz.,

$$p_2^2 = \frac{V_2(\phi)}{T(\phi)}. \tag{33}$$

Moreover, by Rayleigh's theorem, these estimates will in the case of the gravest mode having an assigned number of nodal diameters be too high. The second member of the equation (31), when ϕ has the correct value, is therefore greater than $p_1^2+p_2^2$. Hence in the case supposed the formula

$$p^2 = p_1^2+p_2^2, \tag{34}$$

where p_1 and p_2 are the values of p on the two extreme suppositions of infinite thinness, and infinitely slow rotation, respectively, gives an approximation which is of the nature of a *lower limit* to the true value.*

The property here inferred may be generalised as follows. If the restoring forces which control the vibrations of an elastic system can be separated into two or more groups which affect the potential energy independently, and if the gravest frequency of vibration be found on the assumption that each group acts separately (the inertia being unchanged), then the sum of the squares of the frequencies thus found is less than the square of the greatest frequency which can occur when all the groups act simultaneously; *i.e.*, it is less than the square of the gravest frequency natural to the system.

As a numerical example, suppose that

$$E = 2\times 10^{12}, \qquad \rho = 7{\cdot}8, \qquad a = 60, \qquad h = 1, \qquad \omega = 100\pi,$$

in C.G.S. units. In the gravest mode with two nodal diameters we have, on the hypothesis that $\sigma = \frac{1}{4}$,

$$\frac{p_1^2}{\pi^2} = 2{\cdot}375\times 10^4, \qquad \frac{p_2^2}{\pi^2} = 2{\cdot}16485\times 10^4,$$

from (19) and (22). The formula (34) gives

$$\frac{p}{\pi} = \frac{\sqrt{(p_1^2+p_2^2)}}{\pi} = 213{\cdot}07, \tag{35}$$

as a lower limit to the frequency.

5. We proceed to illustrate the application of Rayleigh's method to the problem.

We assume

$$w = \left(1+\beta\frac{r^2}{a^2}\right)\left(\frac{r}{a}\right)^s \cos s\theta \cos pt, \tag{36}$$

---

* See footnote on p. 272.

where β is a constant to be afterwards chosen so as to make the inferred frequency a minimum. We find

$$V_1 = \tfrac{1}{2}\pi\rho\omega^2 a^2 h\left[\frac{s(s+2)A-s^2B}{s+1}+2\beta\frac{s(s+2)A-s^2B}{s+2}\right.$$
$$\left.+\beta^2\frac{(s^3+4s^2+4s+4)A-s^2(s+2)B)}{(s+2)(s+3)}\right]\cos^2 pt, \qquad (37)$$

$$V_2 = \frac{\pi E h^3}{3(1-\sigma^2)a^2}[8(s+1)\beta^2+2(1-\sigma)\{s^2(s-1)+2\beta s(s^2-1)$$
$$+\beta^2(s+1)(s^2-2)\}]\cos^2 pt, \qquad (38)$$

$$T = \tfrac{1}{2}\pi\rho h a^2\left(\frac{1}{s+1}+\frac{2\beta}{s+2}+\frac{\beta^2}{s+3}\right)p^2\sin^2 pt. \qquad (39)$$

Equating the mean value of V_1+V_2 to that of T, or expressing that the total energy is constant, we have

$$p^2 = \frac{L+2M\beta+N\beta^2}{l+2m\beta+n\beta^2}, \qquad (40)$$

where L, M, N, l, m, n have the values given below. We have to choose β so as to make this fraction a minimum.

The stationary values of p^2 are given by the quadratic

$$(ln-m^2)p^4-(lN+nL-2mM)p^2+LN-M^2 = 0, \qquad (41)$$

and the corresponding values of β by either of the equations

$$\beta = \frac{M-mp^2}{np^2-N} = \frac{L-lp^2}{mp^2-M}. \qquad (42)$$

We write for shortness

$$\zeta^2 = \frac{Eh^2}{\rho a^4}, \qquad (43)$$

so that ζ (like ω) is the reciprocal of a time. The coefficients in (40) have then the values

$$L = \frac{s(s+2)A-s^2B}{s+1}\,\omega^2+\frac{4s^2(s-1)}{3(1+\sigma)}\,\zeta^2, \qquad (44)$$

$$M = \frac{s(s+2)A-s^2B}{s+2}\,\omega^2+\frac{4s(s^2-1)}{3(1+\sigma)}\,\zeta^2, \qquad (45)$$

$$N = \frac{(s^3+4s^2+4s+4)A-s^2(s+2)B}{(s+2)(s+3)}\,\omega^2+\frac{16(s+1)+4(1-\sigma)(s+1)(s^2-2)}{3(1-\sigma^2)}\,\zeta^2, \qquad (46)$$

$$l = \frac{1}{s+1}, \qquad m = \frac{1}{s+2}, \qquad n = \frac{1}{s+3}. \qquad (47)$$

We have specially in view the case of $s = 2$. Adopting the value $\frac{1}{4}$ for σ, we have

$$A = \frac{13}{32}, \qquad B = \frac{7}{32}. \qquad (48)$$

Hence

$$L = \frac{19}{24}\,\omega^2 + \frac{64}{15}\,\zeta^2, \qquad M = \frac{19}{32}\,\omega^2 + \frac{32}{5}\,\zeta^2, \qquad N = \frac{89}{160}\,\omega^2 + \frac{352}{15}\,\zeta^2, \tag{49}$$

$$l = \frac{1}{3}, \qquad m = \frac{1}{4}, \qquad n = \frac{1}{5}. \tag{50}$$

As a partial verification of the method we note that when $\omega = 0$ the quadratic (41) reduces to

$$15\left(\frac{p^2}{\zeta^2}\right)^2 - 19712\frac{p^2}{\zeta^2} + 212992 = 0, \tag{51}$$

the lower root of which is

$$\frac{p^2}{\zeta^2} = 10{\cdot}8955.$$

This makes

$$\frac{p}{\pi} = 1{\cdot}05069\,\zeta. \tag{52}$$

This is higher than the correct value (22), as was inevitable, but the error is less than ½ per cent.

To illustrate the case where the effects of centrifugal force and flexural rigidity are comparable, take the numerical data given at the end of § 4. These make

$$\omega^2 = 9{\cdot}86960 \times 10^4, \quad \zeta^2 = 1{\cdot}97847 \times 10^4, \tag{53}$$

$$L = 1{\cdot}62549 \times 10^6, \qquad M = 1{\cdot}85223 \times 10^6, \qquad N = 5{\cdot}19182 \times 10^6. \tag{54}$$

The quadratic (41) becomes

$$0{\cdot}00417\,p^4 - 1{\cdot}12959 \times 10^5 p^2 + 5{\cdot}00850 \times 10^{10} = 0. \tag{55}$$

The lower root is

$$p^2 = 4{\cdot}50890 \times 10^5, \tag{56}$$

whence, for comparison with (35),

$$\frac{p}{\pi} = 213{\cdot}74. \tag{57}$$

This is probably very close to the true value. In any case the upper and lower limits which we have obtained differ only by about $\frac{1}{5}$ per cent. For practical purposes the approximation (35) would probably be sufficient in all cases, for reasons given above. It is, of course, obtained with much less labour than (57).

The value of β corresponding to (56) is $-0{\cdot}169$.

A Liquid Oxygen Vaporiser.

By Edgar A. Griffiths.

(Communicated by the Tyndall Mining Bequest Committee. Received March 21, 1921.)

Liquid oxygen can be used for a variety of purposes in place of the compressed gas. When the weight of the apparatus for the supply of the oxygen is an important factor, as in the case of breathing apparatus, there is considerable economy in carrying liquid oxygen instead of compressed gas.

For most purposes a steady stream of gas is desired, so means have to be provided for vaporising the liquid at any specified rate from the container. Since the most efficient form of container is the Dewar vacuum vessel, the problem resolves itself to devising a method of controlling the heat input into the liquid when contained in a vessel of this type. An electrical heating coil affords one means of doing this, but the weight of the necessary batteries for supplying the current is a serious drawback when a portable apparatus is desired.

In the course of some experiments on a liquid oxygen apparatus for mine rescue work the author found that it was possible to control the heat conduction into the inner vessel by bringing a small area of the wall of the outer vessel into contact with that of the inner vessel. By varying the area of contact any desired rate of evaporation of the liquid, within certain limits, could be obtained.

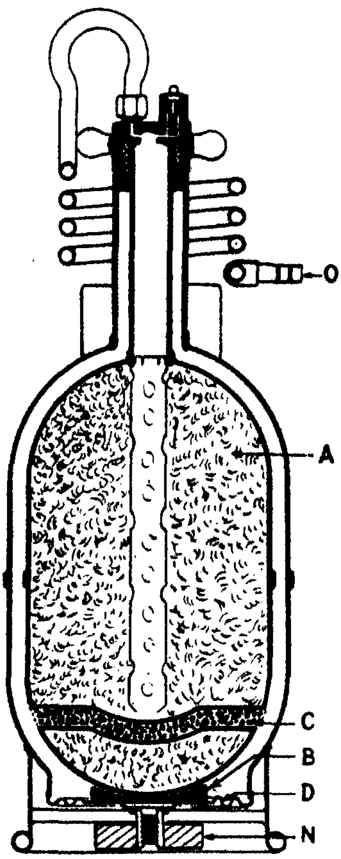


Fig. 1.—A, absorbent material; C, tube through inner vessel containing cocoa-nut charcoal; B, copper block; D, corrugated silver diaphragm; N, screw and nut controlling position of block B.

In practice the apparatus is constructed as shown in fig. 1. The double-walled vessel is made of copper and the bottom of the outer vessel is cut out and replaced by a flexible diaphragm of silver (D). This diaphragm has soldered to its centre a copper block (B) shaped to fit the contour of the inner vessel. A screw and nut (N) is arranged below the diaphragm in order to adjust its position, in or out, thus enabling the block to be brought into contact with the inner vessel at will. Two turns of the adjusting screw are required from "off" to "full on" and this arrangement varies the rate of delivery of the gas from the natural rate of the flask (*i.e.*, about ½ litre per minute) to 12 litres per minute measured at normal temperature and pressure.

The variable rate of gas emission is due to the fact that the copper block

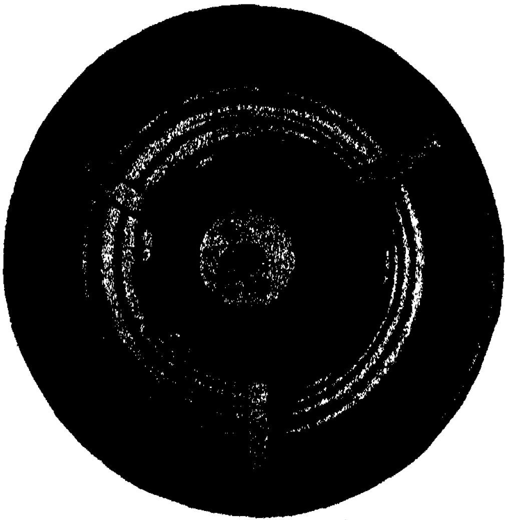


Fig. 2.—Underside of flask showing silver diaphragm and adjusting nut.

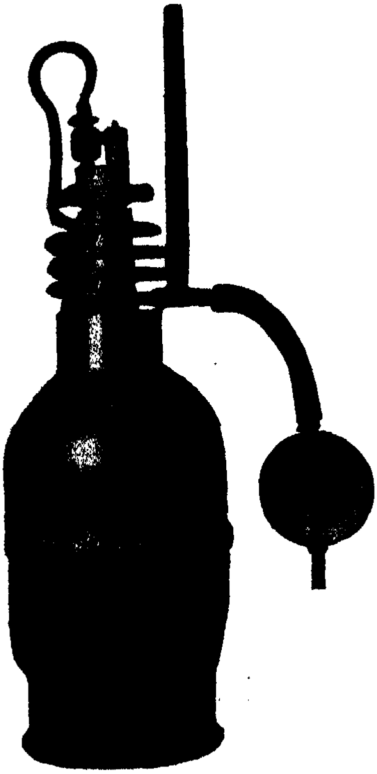


Fig. 3.

does not come instantaneously into contact with the inner vessel over its entire surface but accommodates itself more and more as the screw is released.

The adjustment is perfectly continuous throughout the range, and any desired rate can be obtained within the limits; the usual setting being 3 litres per minute when used for breathing apparatus. The delivery of the gas in response to a change in the adjustment is surprisingly quick and when once set has the convenience of remaining sensibly constant for many hours.

The silver diaphragm and adjusting nut can be seen in fig. 2.

The vacuum vessel used in the experiments had a capacity of about 1½ litres of liquid but was filled with an absorbent packing of asbestos wool to ensure non-spilling when carried. The capacity with absorbent filling was slightly over 1 litre of liquid or 850 litres of gas at N.T.P. The oxygen is delivered through a few turns of copper tube so as to warm up the gas before delivery to the breathing bag.

A simple flow meter is inserted in the circuit (see fig. 3) which indicates at any instant the quantity of gas being supplied. This instrument consists merely of a small vane pivoted so as to swing against the restoring force of a hair spring. The gas enters by way of a wide bore jet and the momentum of the stream causes a deflection of the vane.

The apparatus described was developed in the course of an investigation carried out with the Tyndall Mining Research Studentship of the Royal Society, and the author desires to record his indebtedness to the Tyndall Research Committee for their kind encouragement, also to Sir Joseph Petavel for the use of some of the facilities of the National Physical Laboratory.

A Simple Apparatus for Approximate Harmonic Analysis and for Periodicity Measurements.

By W. Lawrence Balls, M.A., Sc.D. (Cantab.), Late Fellow of St. John's College, Cambridge.

(Communicated by Dr. F. F. Blackman, F.R.S. Received May 2, 1921.)

The apparatus described in this paper was devised in order to test an extensive series of engineering data for the existence therein of a suspected periodicity. In the first instance, we applied the periodogram method* but found that the time needed for adequate examination by this method was prohibitive. Several other devices were tried before we arrived, as a development from a suggestion made by Mr. G. Udny Yule, F.R.S., at the present apparatus, which was found to be not only satisfactory as a method for detecting periodicity but was also useful as an harmonic analyser. The apparatus has the advantage of being simple to construct, and inexpensive. Further, the construction is such that there can be no change of zero during its use, and while the reduction of the observed results to absolute measure requires the employment of corrections, these corrections are mainly based on simple geometry.

---

* 'Edinburgh Mathematical Tracts,' No. 4 (London, 1915), G. A. Carse and G. Shearer, "A Course in Fourier's Analysis and Periodogram Analysis."

In its simplest form the process of periodogram analysis consists in averaging some considerable number, say $2n$, of ordinates taken at a constant interval h along the curved to be analysed, and repeating this averaging for successive series of ordinates starting at $x = 0$, $x = h$, $x = 2h$, $x = 3h$, and so on. If h is exactly equal to any period existing in the curve (or to a multiple of it) and there is no other period, the average ordinate $\bar{y}$ so obtained will show a periodicity of the same amplitude. As h departs from the period, to a greater or less extent, the amplitude of the curve given by $\bar{y}$ falls off. The existence of a number of periods in the curve to be analysed does not alter the general character of the result; in any case $\bar{y}$ shows a maximum amplitude when h is equal to a period. The problem of periodogram analysis, from the mechanical standpoint, thus reduces to the problem of adding or averaging a number of ordinates on a given curve.

Starting with the idea of effecting this addition from a template cut out to represent the required curve, Mr. Yule suggested the employment of a series of coupled springs to effect the addition. The employment of such springs, however, introduces serious constructional difficulties in making them adjustable for various trial periods, from which difficulties the present arrangement escapes. Owing to the appearance of the apparatus, it is commonly described in our laboratory as "The Harp."

Description of Apparatus. A. Essential Parts.

The general appearance of the Harp is shown in plan, and, as a sectional elevation, in fig. 1. A board of triangular shape, 10 feet in length, on a three-point support, is placed with its wider end just overhanging the edge of the table on which it rests. At this wide end of the board are two groups of pulleys (*pp*), each group being spaced at intervals of 25 mm. on a common axle by means of collars with set screws and intermediate washers; meccano parts are quite suitable for the purpose. The innermost pulleys of the two groups are separated by an interval of 50 mm., where the two axles abut in a common support. The four ends of the two axles lie on the arc of a circle struck from the small end of the board, and thus, in view of the small angle subtended by these pulleys (9° of arc), they may be considered effectively as rotating each in a plane which passes through the centre of the circle thus struck, while in number there are effectively nineteen, spaced at equal intervals, but with the centre one missing.

Over each of these pulleys passes a strand of sewing cotton, bearing a weight made of ½-inch lead piping, and weighing about 80 grm., which hangs down outside the edge of the table. The weight on each one of these strings of the Harp is the same, thus keeping all the strings under the same tension, except for accidental variations in the static friction of the pulleys.

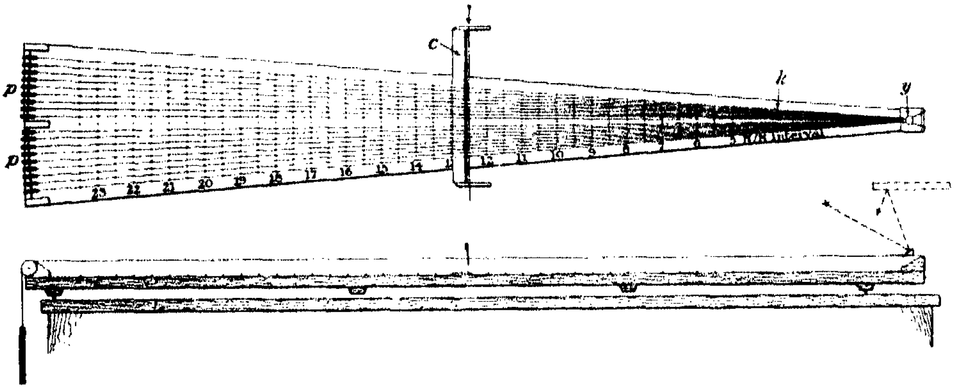


FIG. 1.

The other ends of these strings converge to the neighbourhood of a common meeting point, each one being inserted into a separate hole arranged for it in a yoke of thin sheet brass at the small end of the board. This yoke (*y*, figs. 1, 2A, 2B), is drilled with two rows of nine holes each, which alternate. Each string can thus be brought to and threaded through its appropriate hole, then taken back and knotted on itself in a wider part of the board (*k*), so that, although the eighteen strings are actually in contact laterally, any one may be lifted vertically without disturbing its neighbours.

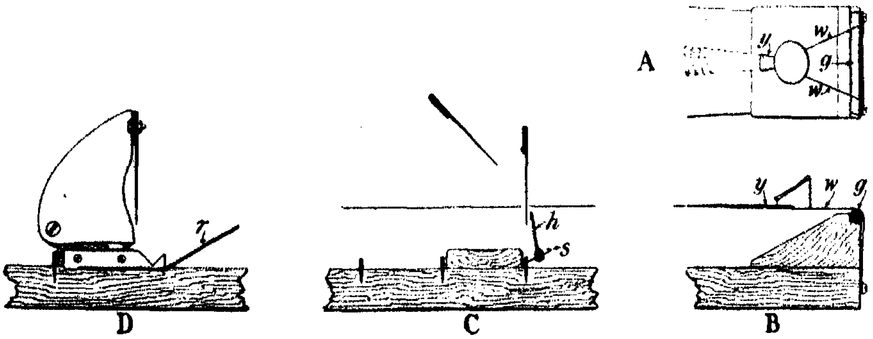


FIG. 2.

The brass yoke (*y*) is soldered firmly to the apex of a triangular loop of steel piano-wire (*w*), 0·01 inch thick (figs. 2A, 2B), the ends of which are fastened by screws to the board after being bent to a right angle round a support of glass rod (g), as shown in the figure. The yoke (y) is thus free to move in a vertical plane under the influence of very small forces, on account of the flexibility of the thin wire, while at the same time it is constrained against horizontal movement or twisting movement. Any such movements of the yoke are magnified and observed by means of a 2-cm. concave mirror, fixed on the yoke in a convenient position by a cork support. A ray of light thrown on the mirror is reflected up to a plane mirror on the ceiling, and thence down to a scale suitably screened by black cloth; this scale may be placed in any convenient position, either by the side of the Harp when the apparatus is being used by a single observer, or elsewhere when two observers are concerned.

Operation of Essential Parts.

The arrangement of strings and yoke thus far described constitutes the essential portion of the apparatus. Considering any one string, this may be regarded as connected to the flexible steel triangle by a hinge at the yoke. Further, the steel triangle is, by virtue of its flexibility, effectively hinged (for movement in a vertical plane) where the steel wire (w) touches the glass rod (g). Under the load of the lead weight hanging from the pulley (p), the yoke hinge lies on a straight line (the weight of the mirror and yoke being negligible) drawn between the upper side of the glass rod and the upper side of the pulley.

Considering all the eighteen strings together, the angle of divergence between the outermost strings is so small that for the purpose of the apparatus they may be considered as effectively parallel to one another at any point along their length. If the angle between the outermost string and the centre line is θ, the outermost string exerts an effective component pull on the yoke of $f \cos \theta$ instead of the pull f exerted by the central string. As θ is only 4°, $\cos \theta$ is 0·995, and the difference of one-half of 1 per cent. is quite negligible. Further, at any such point any number of the strings may be raised above the zero position to any extent required, increase in length of the working portion of the string being permitted by the rise of the weight, the pulley turning to permit this rise, and the tension being unaffected.

If, now, all the eighteen strings are raised to the same extent at some place along the board (*e.g.*, by placing a straight edge underneath the strings), the wire with its yoke and mirror will bend until a new straight line position is assumed, so that the yoke hinge lies in the plane which contains the top of the glass rod and the straight edge respectively. If, on the other hand, only half the strings are thus raised, and this irrespective of the individual strings chosen, the yoke hinge will rise by only half the previous amount, and the movement of the spot of light reflected from the mirror to the reading scale will also be halved. This strict proportionality will, of course, only hold good so long as the deflection of the mirror is not great enough to be affected by the stiffness of the wire (*w*), but in the actual operation of the apparatus, this deflection cannot exceed 4° of arc, and is not usually more than 2°.

Similarly, for any number of strings lifted from the neutral position to a given height above it, the position of the mirror will change in proportion to the number of strings lifted.

Again, for any one string lifted, the deflection of the mirror will be proportionate to the amount of the lift, and similarly, again, for any number of strings lifted to varying extents. For small deflections, *i.e.*, those lying within 1° of arc, the deflection of the mirror is practically in direct proportion to the average deflection of the strings above their zero position. For larger deflections of the strings, the movements of the mirror fall off somewhat, but not sufficiently to hamper the use of the appliance.

Having shown the means whereby the spot of light on the observing scale is enabled to indicate the average position of all the eighteen strings, it remains to describe the accessory parts by which the instrument is made convenient for actual use.

Description of Apparatus. B. Accessory Parts.

The triangular board is scaled off to make a series of trial periods by utilising the divergence of the strings. Underneath each string a straight line is ruled, radiating from the common centre below the yoke, and on each of these lines, at regular intervals, a series of steel pegs is driven, as indicated in fig. 1. These pegs thus form a series of rows across the board, beginning with a row whose members are 5 mm. apart, and increasing these intervals by ½-mm. steps up to 23½ mm. near the pulleys. These pegs serve a double purpose: in the first place they act as sounders in combination with the click arrangement, to be described shortly, while in the second place they serve to register the positions at which the template may be passed under the strings.

The template is passed under the strings by attaching it to a frame which slides on a carriage (*c*, fig. 1). This carriage is not fixed to the board, but can be placed thereon at any one of the positions defined by the rows of steel pegs. In fig. 1, it is represented in position for testing a trial period of 12½ mm., and the line of movement of the template is indicated by an arrow.

This carriage consists of two parts, namely, a comb, and also the guide in which the template slide moves. The comb consists of a large number of steel needles, about 8 cm. in length, obtained by breaking the heads off ordinary drapers' round-headed bonnet pins, which are then set up in a row and soldered on to a strip of $\frac{1}{8}$-inch brass to form a comb with about eight teeth per centimetre, and of sufficient length to cover the greatest spread of the strings at the wide end of the board. The object of this comb is to prevent any strings from being pushed sideways out of their proper position by friction with the template passing below them. The effect of its use is necessarily to disturb the exact identity of the intervals between the strings, but the error caused thereby is not serious; thus in fig. 3, the comb and the strings are drawn diagrammatically to scale for an interval between the strings of 10 mm. In order to enable this comb to be placed rapidly over the strings at any position along the board, it is mounted on a hinged frame shown in fig. 2D, in such a way that when the comb is lifted up, its points swing clear of the strings; in this position, the whole carriage can be transferred to any other part of the board and placed between two other rows of pegs, whereupon, on lowering the comb, the points of the needles deflect the strings slightly to the right or left if necessary. The friction of the strings on the teeth of the comb introduces a slight error in the nominal equality of tension obtaining between all the strings, but this again is not serious.

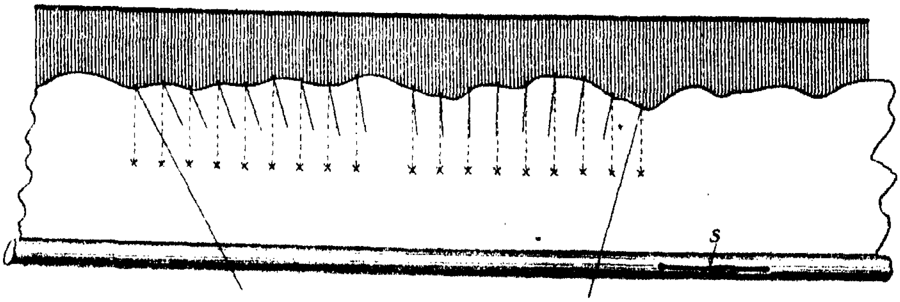


FIG. 3.

The comb also serves to define exactly the position of the edge of the template, since the template leans up against the front of the comb while sliding across the carriage. In fig. 2D, an end elevation of the carriage, is shown one of the V-shaped notches fixed on the end of the wooden base of the carriage, in which the template holder slides. This holder may be of any convenient form, such as a straight bar of metal to which is fixed a wing of tin plate (*h* in fig. 2C), against which the cardboard template can be clipped. The ends of the slider may be equipped with spools on which to coil up portions of the template not under examination. In addition, this sliding bar bears a clicking spring (figs. 2c and 3, *s*), so shaped as to make contact with each one in any row of eighteen pins, already described as fastened to the base board, and to emit an audible click on leaving each pin. This slide should be tested in the Harp with a straightedge as a template, to detect any periodicity in the slide.

When the carriage is to be moved to another part of the board for examination of a different trial period, the slide, with its attached template, is turned over (past its dead centre on the strings) until it rests on two projecting arms (fig. 2, *r*) so as not to foul the under side of the strings while the carriage is being moved.

Operation of Accessory Parts.

A smooth template having been prepared, it is fixed on the slide with its lower edge in contact with the bar thereof. The scale of the template being such that there is reason to expect the occurrence of a period of length lying between 5 and 23 mm., the carriage is then placed on the board, with its base fitting between the 5 and 5½ mm. rows of steel pegs. The slide with its attached template, is then passed under the strings, and placed with its bar resting on the V-notches of the carriage. The template is then rotated in the notches until it makes contact with the strings, and, passing the dead centre, is stopped by contact with the comb as well, where the strings are 5 mm. apart. The slide is then pushed along slowly by hand, thereby causing the strings to rise and fall as they follow the outline of the curve. At every fifth millimetre of this traverse the click is heard. Meanwhile, the operator or another observer is watching the movement of the spot of light. If a period 5 mm. in length actually exists in the template as plotted, the spot of light will oscillate in synchronism with the clicks.

If no such synchronism is noticed between the clicks which are heard and the light movements which are seen, the template is turned down on to its rests, the comb swung up and back, the carriage lifted and replaced at the next interval, and the operation repeated to test for the existence of a period at 5½ mm. Similarly, all periods are tested from one end of the board to the other. The full search over this range, from 5 mm. to 23 mm., by ½-mm. steps, does not require more than a quarter of an hour to complete.

A trial period having thus been found at which the click and the movement of the light are synchronised, the length of the period is at once shown by the position of the template on the board. The amplitude and the form of the periodic curve can be measured at this setting as follows, preferably by two observers co-operating:—

The slide carrying the template has a millimetre scale marked on it, by means of which the position of the slide is advanced by small and equal successive steps, while the position of the spot of light is recorded for each step on millimetre-squared paper. The height of each ordinate, representing the average position of the strings at each step, may either be chosen arbitrarily, and subsequently corrected, or may be chosen in agreement with a scale previously prepared to show the actual deflection of the spot of light for a given mean movement of the strings at this particular spot on the board; in either way, a value for the amplitude of the periodic curve may be obtained, which is usually within 3 per cent. of the true value.

Preparation of Templates.

One or two simple devices which have been found useful are possibly worth mentioning. The paper employed for the making of templates requires to be in long strips not more than 3 inches in width, and the most convenient way of obtaining this is to have three rolls of paper, such as those from a Burrough's adding machine, pasted together to form a flexible cardboard.

This paper is, of course, not squared for plotting, but curves drawn on squared paper can easily be transferred to it by tracing through typewriter carbon paper.

The scale of the curve should be so chosen that in the forward direction there should be no slope exceeding about 45°, since a steeper slope than this will tend to wedge the strings against the comb, and prevent them from rising and falling smoothly.

If the template, after being traced, is merely cut out with scissors and left thus, the rough edge will tend to make the strings vibrate, and cause an inconvenient flicker of the spot of light. This can be obviated by waxing the edge of the template after cutting, preferably with a hard wax, such as that used for making phonograph records, and finally running a hot iron along the edge, to smooth the wax down to the edge of the paper.

Bearing in mind the necessity for avoiding steep slopes in the curves, the scales to which the curves are plotted should be so chosen as to give a total range of about 3 cm. between the highest and lowest portions, for although the instrument as here described can accommodate curves with an amplitude of 5 cm., it is obvious that the advantage from greater movement of the spot of light will tend to be counter-balanced by excessive distortion of the strings. Again, if any preliminary indication exists as to the probable length of the period, the scale should be so chosen as to make the period fall in the neighbourhood of the middle of the board, for the same reason.

In the event of the search requiring to be continued beyond the limits of 5 mm. and 23 mm. on the scale chosen, the curve must, of course, be re-drawn, and a new template cut to a larger or smaller magnification. While the dimensions here given for the instrument are convenient in an ordinary laboratory, there is no reason why they should not be exceeded by increasing the length of the board, so as to give a greater range of search on any one template, since this range at present is such as necessitates the templates being cut at four-fold magnifications and reductions. In this event, it should be borne in mind that the sensitivity of the instrument decreases as the distance of the template from the mirror increases.

Uses of the Apparatus.

We have already described the method of using the apparatus for determining primarily the length but also the amplitude and form of any periodicity existing in a curve. It will be seen that, by suitable choice of templates, the same procedure can be applied to Harmonic Analysis, the single curve being plotted repeatedly on the template until a sufficient length is available to pass through the instrument, the length of each period being made about 40 mm. In this way, the first four or five harmonics can be detected on the same template, and their amplitude measured by the same procedure as before.

Another occasional use for the apparatus is in the smoothing of curves, which can be passed in template form under the strings at any desired setting, and with any number of strings in use, according to the number of points which we require to include in the smoothing, while a second observer plots the readings of the spot of light as the template moves forward.

Summary.

An account is given of a simple and inexpensive means for constructing and using an apparatus by which periodicities can be determined, harmonic analysis effected, and curves smoothed.

The results obtained (both with respect to length and to amplitude) are only approximately correct, but the error involved need not exceed 3 per cent.

The outstanding advantage of the apparatus is the speed with which determinations may be made. Thus, in the case of determinations of periodicity, some fifty trial periods can be examined in less time than is required for the computation of a single trial period under the periodogram arithmetical method.

I have pleasure in acknowledging the assistance of Mr. G. Udny Yule, F.R.S., while the extensive use of the Harp, which enabled its utility to be fairly judged, has been in the hands of my assistants, Miss E. Lanigan and Mr. H. Hancock, in the Experimental Department of the Fine Cotton Spinners' Association at Manchester.

Multenions and Differential Invariants.

By Alex. McAulay, M.A., Professor of Mathematics, University of Tasmania.

(Communicated by Mr. W. B. Hardy, Sec.R.S.—Received May 13, 1920.)

§ 1. *Preliminary Explanations.*—Since multenions appear almost to have been designed by Nature to serve as an algebra dealing with such matters as differential invariants and relativity, I have thought it desirable at the present moment to present a summary of their properties. Towards the end I propose to enter into more detail in applying them to differential invariants, and to point out some of the special properties of multenions when n, the vector complexity of the system, has the particular value 4.

Clifford first, I believe, applied the subject to elliptic space of three dimensions. His bi-quaternion $q+\omega r$, where $\omega^2 = 1$, is the general member of a sub-algebra of multenions when $n = 4$. In the form of multenions suitable to deal with real questions of Relativity Hamilton's bi-quaternion $q+r\surd(-1)$ presents itself in a precisely similar manner, and of this I shall have more to say towards the end of the paper. Joly's 'Manual of Quaternions' has its last chapter devoted wholly to multenions. I believe the most elaborate attempt at development is my own paper in 'Proc. Roy. Soc., Edin.,' 1907–8, p. 503.

The notation used below is based as closely as possible on quaternion notation. It is practically identical with Joly's so far as the latter goes. It differs in several respects from the notation of my own former paper. The changes are all in the direction of simplicity, and back towards quaternion rules. The same applies to changes in terminology. These changes, which I believe are all for the better, are in the main due to friendly advice received from Prof. Knott in connection with the former paper.

Law I. Besides the unit scalar 1 there are given n generators, which will generally be regarded as representing n unit mutually orthogonal vectors in a

Euclidian space of n dimensions, $i_1, i_2, \dots i_n$. These are such that all the ordinary laws of algebra except the commutative law of multiplication apply to the general integral expression q, called a multenion, where

$$q = \Sigma\,(x_0 + x_1 i_1 + x_2 i_2 + \dots)(x_0' + x_1' i_1 + x_2' i_2 + \dots)\dots, \tag{1}$$

where every x is a scalar. Since any x may be negative we understand this statement to include the ordinary algebraic uses of the signs $+$ and $-$; but we do not understand that any assertion has been made about the use of the sign $\div$ [$q \div r$ can later be understood to be an alternative for qr^{-1}].

Law II. Scalars are commutative not only with each other but with each of the n primitive unit vectors. [Thus scalars are commutative with multenions.]

Law III. [i_1^2 means $i_1 i_1$, etc.]

$$-1 = i_1^2 = i_2^2 = \dots = i_n^2.$$

Law IV. In a product, called a primitive vectorium, $i_a i_b \dots$, of primitive unit vectors, if two adjacent *different* vectors be interchanged, the product merely changes sign.

Such are the assumptions. The rest is development, aided of course by suitable terminology and symbolisation.

Def. 1. A primitive unit means any one of the following: (1) the scalar 1, (2) any primitive unit vector i_a, (3) any other primitive vectorium $i_a i_b \dots$. [Strictly (3) includes (2) and in a sense (1).]

Def. 2. An expression of the form $\rho = \sum_{c=1}^{n} x_c i_c$ where each x is a scalar, is called a vector, or a V_1 or a ${}_n\mathrm{V}_1$,

Def. 3. Any multenion is obviously capable of expression in the form $\Sigma x i_a i_b \dots i_k$ where $i_a, i_b, \dots i_k$ are m different primitive vectors and m ranges from 0 to n, m being given $\mathrm{V}_m q$ means the sum of all parts of the multenion for that value of m. Thus

$$q = \mathrm{V}_0 q + \mathrm{V}_1 q + \mathrm{V}_2 q + \dots + \mathrm{V}_n q.$$

$\mathrm{V}_a q$ is said to be the part of q of homogeneity a and is called a V_a or a ${}_n\mathrm{V}_a$ or a homogeneous multenion or a hyper vector. [These alternatives are intended to invite suggestions for a settlement of the name. I should have preferred "an a-vector" but this clashes with relativity meanings of 4-vector and 6-vector. In multenions with $n = 4$, there are two kinds of 4-vector ${}_4\mathrm{V}_1$ and ${}_4\mathrm{V}_3$.]

$\mathrm{V}_0 q$ will generally in words be called the scalar part of q but in the present connection we must not use $\mathrm{S}q$ as an alternative for $\mathrm{V}_0 q$. In applications we very decidedly need to be able, in one system of multenions ($n = 4$) to make no change in the quaternion notation but yet to use both the quaternion notation and the full multenion notation as applying to one symbol q. Thus

$$Sq = V_0 q + V_4 q, \qquad Vq = V_2 q.$$

Endless confusion would result from allowing V_0 to be represented by S.

Def. 4. If $\alpha_1, \alpha_2, \ldots \alpha_a$ are any a-vectors then $V_a \alpha_1 \alpha_2 \ldots \alpha_a$ is called a vectorium. [Hence we have already called such an expression as $\iota_1 \iota_2 \iota_3$ a primitive vectorium.]

Def. 5. A linear multenion function of a multenion is called a multenion linity. Similarly for a vector linity or a ${}_n V_a$ linity, etc. An extended vector linity is a special kind of multenion linity, Φ, defined from a given vector linity, ϕ. Its fundamental property is this: If $\alpha_1, \alpha_2, \ldots \alpha_a$ are any a-vectors whatever, a itself ranging from 0 to n, then

$$\Phi V_a \alpha_1 \alpha_2 \ldots \alpha_a = V_a \phi\alpha_1 \phi\alpha_2 \ldots \phi\alpha_a. \tag{2}$$

[It is not obvious that this remark covers a possibility in general. The definition, which will be modified below, is given here as a hint of the importance of vectoriums and extended linities in our subject.]

Usual Notations.

Scalars (V_0): a, b, c, g, h, k, l, m, n, t (= time), x, y, z, J (= $\sqrt{[-1]}$), θ (= arc), π.

General multenions: p, q, r, s.

Hyper vectors (homog. mult.) and vectoriums:

Homogeneities	V_a a	V_b b	V_c c	V_n n	V_2 2
Hyper vectors	u	v	w	$x\dot{v}$	ω
Vectoriums...........................	$\dot{\alpha}$	$\dot{\epsilon}$	$\dot{\eta}$	$x\dot{\upsilon}$	$\dot{\omega}$

(These conventions should be noted and remembered. They immensely simplify the reading of formulæ by their property of minimising number of suffixes, etc.)

Vectors (V_1): α, β. γ, δ, ϵ, ζ, ξ, η, λ, μ, ν, ρ, σ, τ.

Linear functions; ξ, η, λ, μ, ν, ϖ, υ, ϕ, χ, ψ.

Vectoriums: $\dot{\alpha}$, $\dot{\epsilon}$, $\dot{\eta}$, $\dot{\iota}$, $\dot{o}$, $\dot{\rho}$, $\dot{\upsilon}$, $\dot{\omega}$.

Differential symbols: δ, Δ, H, L, d, $d\rho$, $d\dot{\rho}_b$, $d\varsigma_a$, $d\alpha$, db.

Selective linities, etc.: K, P, Q, R, S, T, U, V.

Def. 6. Complexes of vectors, hypervectors, multenions, etc., are sufficiently illustrated by multenions. If $q_1, q_2, \ldots, q_a$ are a given multenions, then all multenions of the form $\sum_{c=1}^{a} x_c q_c$ are said to form a complex. If $q_1, q_2, \ldots q_a$

can be made linearly dependent on b and no fewer, then b is called the complexity of the complex. Thus the complexity of the complex of all vectors is n, of all multenions it is 2^n, of all multenions of homogeneity a it is ${}_nC_a$.

Def. 7. A unit vector or a unit vectorium is one, $\acute{a}$, such that $\acute{a}^2 = \pm 1$. [The square of every vectorium is a scalar, but u^2 is a scalar in general only in the four cases when $a = 0, 1, n-1$ or n.] Thus all the primitive vectoriums are unit vectoriums.

Def. 8. Two vectors, α, β, are said to be normal to one another when $V_0\alpha\beta = 0$. Two vector complexes are said to be completely normal to one another when every vector of the one complex is normal to every vector of the other.

Every one of the symbols V_a is a multenion linity. They are all commutative with one another. They are all idempotent, that is $V_a^2 = V_a$, and they all satisfy the equation $V_aV_b = 0$ when $a \neq b$.

There is another system of multenion linities of great utility and simplicity. They are 2^n in number. They are all commutative with each other and with each of the $n+1$ linities V_a. Let P_a be that definite linity of a general multenion, q, which simply reverses the sign of every i_a which occurs in the expression (1) for q. From the well-known properties of linities a product of the P's such as $P_1P_2P_3$ simply reverses the sign of every i_1, i_2, and i_3. These linities are (1) unipotent, that is, if

$$\phi = P_aP_b \dots P_k, \tag{3}$$

then $\phi^2 = 1$; (2) they are proscriptive, by which is meant that

$$\phi(qr) = \phi q\, \phi r, \tag{4}$$

where q and r are any two multenions. It is convenient to use P for the product $P_1P_2 \dots P_n$ of all the linities P_a, so that P reverses the sign of every primitive vector.

K is defined, consistently with its quaternion use, in the following manner. Kq is a linity of a general multenion, q, which reverses the sequence of the primitive unit vectors in every primitive vectorium $\acute{\iota}$, of q, *and also* reverses the sign of every primitive unit vector. Thus

$$K(i_1i_2i_3) = (-i_3)(-i_2)(-i_1) = i_1i_2i_3.$$

K is (1) unipotent, that is $K^2 = 1$; (2) it is commutative with every V and every P; (3) it is retroscriptive, that is $K(qr) = KrKq$. [To prove (3) first prove it when q and r are two primitive vectoriums and then generalise. Such a method of proof is very often applicable and may be referred to as proof by reduction to primitive units.]

But from these properties not only is

$$K(qr) = Kr \,.\, Kq, \tag{5}$$

but if ϕ be given by (4), $K\phi\,(=\phi K)$ is also retroscriptive. More generally, if ϕ, ψ, χ, ..., are any number of multonion linities, every one of which is either proscriptive or retroscriptive, then also is the product, ϕ, ψ, χ, ..., proscriptive or retroscriptive; it is proscriptive when the number of ϕ, ψ, χ, ..., which are retroscriptive is even, and it is retroscriptive when that number is odd.

The special retroscriptive PK (= KP) is of sufficient importance to have a symbol appropriated to it. We shall call it Q. Thus the product of any two of the three P, Q, K is equal to the third.

A caution to users of present methods in relativity and hyperbolic geometry seems desirable. It relates to the discussion of real questions by aid of the imagery

$$J = \sqrt{(-1)}. \tag{6}$$

Let us illustrate by anticipating our mode of treatment of real questions in relativity. We begin with the multenion system based on the primitive generators i_1, i_2, i_3, i_4. Then we put

$$i_4 = J\iota, \tag{7}$$

and drop out of most of our subsequent work all reference to the original i_4 and J_2, using ι instead with its property $\iota^2 = 1$, as opposed to $i_3^2 = -1$. Questions in relativity are then treated by aid of the multiplicative combinations of such expressions as

$$xi_1 + yi_2 + zi_3 + ct\iota, \tag{8}$$

whereby, so long as all the scalars thereby introduced are real, we shall be dealing with real questions in relativity. The general multenion thus constructed contains 16 and not 32 independent real scalars. ι is for practical relativity purposes real. We shall call such a multenion system semi-real.

Now return to the meaning above assigned to K.

$$K(i_1 i_2 i_3 i_4) = i_1 i_2 i_3 i_4,$$

and also

$$K(i_1 i_2 i_3 \iota) = i_1 i_2 i_3 \iota,$$

but whereas

$$K(i_1 i_2 i_3 i_4) = (i_1 i_2 i_3 i_4)^{-1}, \tag{9}$$

it is not true that the corresponding equation holds when we replace i_4 by ι. I have therefore refrained from making the definition of K depend on such equations as (9). It is, nevertheless, a perfectly easy theorem to prove that

$$K\acute{\imath} = \acute{\imath}^{-1}, \tag{10}$$

when $\acute{\imath}$ is any product of our *original* generators i_1, i_2, ... i_n. This ceases to be true when ι takes the place of i_4, simply because J^{-1} is $-J$ and not $+J$.

It may be noted, however, that there is always a reproscriptive $K\phi = R$ to be found, which is such that for every semi-real vectorium $\acute{\imath}$, $R\acute{\imath} = \acute{\imath}^{-1}$.

Thus, suppose we replace $i_{a+1}, \dots i_n$ by $l_{a+1} \dots l_n$

where $$i_c = \mathrm{J}l_c \text{ and therefore } l_c^2 = 1, \qquad c = a+1, a+2, \dots n. \tag{11}$$

Then $$\mathrm{R} = \mathrm{KP}_{a+1}\mathrm{P}_{a+2}\dots\mathrm{P}_n = \mathrm{QP}_1\mathrm{P}_2\dots\mathrm{P}_a, \tag{12}$$

is easily shown to have the required characteristic.

We now obviously have

$$\mathrm{P} = \mathrm{P}_1\mathrm{P}_2\dots\mathrm{P}_n, \tag{13}$$

$$\mathrm{P}u = (-)^a u, \qquad \mathrm{Q}u = (-)^{\frac{1}{2}a(a-1)}u, \qquad \mathrm{K}u = (-)^{\frac{1}{2}a(a+1)}u. \tag{14}$$

Further if we put

$$q_c = \mathrm{V}_c q, \qquad c = 0, 1, \dots n, \tag{15}$$

$$\left.\begin{aligned} \mathrm{P}q &= q_0 - q_1 + q_2 - q_3 + \dots \\ \mathrm{P}_{(0)}q &= \tfrac{1}{2}(1+\mathrm{P})q = q_0 + q_2 + q_4 + \dots \\ \mathrm{P}_{(1)}q &= \tfrac{1}{2}(1-\mathrm{P})q = q_1 + q_3 + q_5 + \dots \end{aligned}\right\}, \tag{16}$$

$$\mathrm{Q}q = q_0 + q_1 - q_2 - q_3 + q_4 + q_5 - \dots, \tag{17}$$

$$\mathrm{K}q = q_0 - q_1 - q_2 + q_3 + q_4 - q_5 - \dots, \tag{18}$$

Clearly, we might introduce Q_0, Q_1, and K_0, K_1, corresponding to $\mathrm{P}_{(0)}$ and $\mathrm{P}_{(1)}$.

All the linities recently considered (V_c, ϕ, $\mathrm{K}\phi$) may be referred to as the capital linities.

§ 2. *Theorem of Independence and Allied Algebraic Features.*— If $q_1q_2 = -q_2q_1$, then q_1 and q_2 are said to be anti-commutative. In the following enunciation, a product of a different multenions from among $q_1, q_2, \dots q_m$ is said to be odd-formed or even-formed, according as a is odd or even.

If $q_1, q_2, \dots q_m$ be m multenions such that $q_1^2, q_2^2, \dots q_m^2$ are all scalars differing from zero, and each pair is anti-commutative; then, if m is even, the 2^m multiplicative combinations are linearly independent; and, if m is odd, they are independent unless the product $q_1q_2\dots q_m$ of all of them is a scalar, and, if it is a scalar, 2^{m-1} and only 2^{m-1} of the combinations are independent, and these 2^{m-1} independent combinations may be taken as the odd-formed combinations, or instead as the even-formed combinations, or instead as the combinations from which one assigned multenion, such as q_m, is absent. The theorem is true of linities as well as of multenions.

Thus, if n is even, a general multenion cannot be expressed by fewer than 2^n independent scalars. In the case of $n = 4$, it cannot be expressed by fewer than sixteen independent scalars. In the case of $n = 3$, it *can* be expressed by four independent scalars, the same number as in the case of $n = 2$. [But to express it thus, we have to impose the condition that $i_1i_2i_3$ is a scalar such as -1. This is practically Hamilton's procedure in Quaternions.] Either of these cases might be regarded as the case of

Quaternions. I prefer to view Quaternions from another side, namely, as forming the subalgebra of even homogeneities for the case $n = 3$.

In general multenion theory, it is convenient deliberately to prescribe for the case n odd that the 2^n primitive vectoriums shall be independent. This is equivalent to saying that the product $\iota = i_1 i_2 \dots i_n$ of all the primitive vectors is *defined* as independent of the scalar 1, even when n is odd.

Whatever be the value of n, multenions with even homogeneities only form a subalgebra with 2^{n-1} units. When n is odd, this subalgebra can always be identified with the next lower multenion system.

When n is even, the number of independent scalars, 2^n, is a perfect square. In this case the algebra is the equivalent of an algebra of linities with the same number. The characteristic equation, therefore, satisfied by a multenion, may be found. When n is even, the degree of the characteristic equation is $2^{\frac{1}{2}n}$. When n is odd, the degree is $2^{\frac{1}{2}(n+1)}$.

In the case of $n = 4$ or 3, the characteristic function may be written in a variety of allied forms, of which the two most important are

$$chf_q(x) \equiv (q-x)\,\mathrm{K}\,(q-x)\,.\,(2\mathrm{V}_0-1)\,[(q-x)\,\mathrm{K}\,(q-x)] \tag{1}$$

$$\equiv (q-x)\,\mathrm{Q}\,(q-x)\,.\,(2\mathrm{V}_0-1)\,[(q-x)\,\mathrm{Q}\,(q-x)]. \tag{2}$$

The coefficient of every power of x is a scalar which it is not difficult to write down in full. Suppose these scalars are m, m', m'', m''', thus

$$chf_q(x) = x^4 - m'''x^3 + m''x^2 - m'x + m. \tag{3}$$

This defines the scalars m. We now have the identity

$$q^4 - m'''q^3 + m''q^2 - m'q + m \equiv 0. \tag{4}$$

In the case of the semi-real system suitable to relativity, $chf_q(0)$ is positive or zero, never negative. The positive value of its fourth root is the natural generalisation of the quaternion tensor $+\sqrt{(q\mathrm{K}q)}$. I have been uncertain whether to denote this positive scalar by $\mathrm{T}q$, thus

$$\left.\begin{aligned} \mathrm{T}q &= +\{q\mathrm{K}q\,.\,(2\mathrm{V}_0-1)\,.\,(q\mathrm{K}q)\}^{\frac{1}{4}} \\ &= +\{q\mathrm{Q}q\,.\,(2\mathrm{V}_0-1)\,.\,(q\mathrm{Q}q)\}^{\frac{1}{4}} \end{aligned}\right\}. \tag{5}$$

An objection to doing so is that this is not Hamilton's meaning in the case of his biquaternion $q+\mathrm{J}r$. His tensor is then imaginary, namely, one of the values of $[(q+\mathrm{J}r)\,\mathrm{K}\,(q+\mathrm{J}r)]^{\frac{1}{2}}$. In our form below, this amounts to saying that the tensor consists of a V_0 part + a V_4 part.

It suffices for our purpose to consider that q only has a tensor when $q\mathrm{K}q$ is a scalar (V_0). For our semi-real relativity case this must be a real scalar, and in that case Hamilton's meaning, and the meaning given by (5), coalesce. [If qKq is a scalar, then

$$qKq = (2V_0-1) \,.\, qKq,$$

and therefore Tq above $= \{(qKq)^2\}^{\frac{1}{2}}$.]

The general value of q for which qKq is a scalar is

$$q = xe^p \text{ where } (1+K)p = 0. \tag{6}$$

The following theorems will now be easily proved by any reader slightly familiar with quaternions.

$$q = \Sigma i V_0 i^{-1} q = \Sigma i^{-1} V_0 i q, \tag{7}$$

$$\rho = \sum_{c=1}^{n} i_c V_0 i_c^{-1} \rho = \sum_{c=1}^{n} i_c^{-1} V_0 i_c \rho, \tag{8}$$

$$V_0(qV_c r) = V_0(rV_c q) = V_0(V_c q V_c r), \tag{9}$$

$$V_0 qr = V_0 rq, \tag{10}$$

whence the cyclic form

$$V_0 q_1 q_2 \ldots q_{m-1} q_m = V_0 q_m q_1 q_2 \ldots q_{m-1},$$

$$V_0(qRr) = V_0(rRq), \tag{11}$$

$$V_0(qKRr) = V_0(rKRq). \tag{12}$$

Here R is any capital retroscriptive and KR any capital proscriptive. Thus, for the capital linities, not only are they all commutative, but they are all self-conjugate. Conjugacy is defined and tested as in quaternions. Formally, if ϕ is a given multenion linity, its conjugate, ϕ', is the unique linity for which

$$V_0 q\phi r = V_0 r\phi' q, \tag{13}$$

q and r being arbitrary. When $\phi = \phi'$, ϕ is self-conjugate. More generally, for any ϕ, $\frac{1}{2}(\phi+\phi')$ is the self-conjugate part and $\frac{1}{2}(\phi-\phi')$ is the skew part. More general kinds of conjugate present themselves occasionally. For instance, let R be any one of the capital retroscriptives. Q_R, the R-conjugate of ϕ, may be defined by $V_0 qR\phi r = V_0 rR\phi_R q$. With these meanings of the conjugate, ϕ', and the R-conjugate, ϕ_R of ϕ, we have

$$\left.\begin{array}{lll} (\phi+\psi)' = \phi'+\psi', & (\phi\psi)' = \psi'\phi', & (\phi')' = \phi \\ (\phi+\psi)_R = \phi_R+\psi_R, & (\phi\psi)_R = \psi_R\phi_R, & (\phi_R)_R = \phi \\ \phi_R = R\phi' R, & \phi' = R\phi_R R & \end{array}\right\}. \tag{14}$$

The general meaning of a conjugate, ϕ_c, of ϕ may be taken to be that it is a unipotent retroscriptive linity of the linity ϕ. Thus ϕ' is a linity function of the linity ϕ: (1) unipotent because $(\phi')' = \phi$; (2) retroscriptive because $(\phi\psi)' = \psi'\phi'$; (3) linear because $(\phi+\psi)' = \phi'+\psi'$. I have before me in MS. a *full analysis* of the properties of such a conjugate, but must not give it here. [Rq stands to q as ϕ' stands to ϕ.] The capital linities are self-R-conjugate as well as self-conjugate. If $\phi q = pqr$, then

$$\phi' q = rqp, \qquad \phi_{\mathrm{R}} q = \mathrm{R}p \,.\, q \,.\, \mathrm{R}r, \tag{15}$$

and similarly if $\phi q = p_1 q r_1 + p_2 q r_2 + \dots$.

Frequently, as in proving the above, we require

$$\text{if} \qquad \mathrm{V}_0 pr = \mathrm{V}_0 qr, \text{ then } p = q, \tag{16}$$

r being arbitrary, and again

$$\text{if} \qquad \mathrm{V}_0 \alpha\rho = \mathrm{V}_0 \beta\rho, \text{ then } \alpha = \beta, \tag{17}$$

ρ being arbitrary, and similarly for a ${}_n\mathrm{V}_a$ instead of a general multenion or a vector.

When in a semi-real system R is chosen so that $\mathrm{R}\iota = \iota^{-1}$ for every ι, we have that if $q = \Sigma x\iota$ and $q' = \Sigma x'\iota$, then

$$\mathrm{V}_0 q \mathrm{R} q' = \Sigma x x',$$

and in particular $\qquad \mathrm{V}_0 q \mathrm{R} q = \Sigma x^2.$

§ 3. *The Commutation Theorem.*—If ι_a is of homogeneity a, and ι_b of homogeneity b, and if there are exactly c primitive vectors common to the a vectors of ι_a and the b vectors of ι_b it is easy to see that $\iota_a\iota_b$ is of homogeneity $a+b-2c$, and that $\iota_a\iota_b = (-)^{ab-c}\iota_b\iota_a$. Hence

$$\mathrm{V}_{a+b-2x} uv = (-)^{ab-x}\, \mathrm{V}_{a+b-2x} vu. \tag{1}$$

[Obviously true when $u = \iota_a$, $v = \iota_b$ as we see by considering separately the two cases $x = c$, $x \neq c$. It is therefore true for u and v. This is a case of proof by reduction to primitive units.]

(1) may be called a first form of the commutation theorem as it shows us a first effect of commuting u and v in uv. (1) at once leads to what may be called the full form

$$\left.\begin{aligned} uv &= (\mathrm{V}_{a+b} + \mathrm{V}_{a+b-2} + \dots + \mathrm{V}_{a-b})\, uv \\ &= (-)^{ab}(\mathrm{V}_{a+b} - \mathrm{V}_{a+b-2} + \dots + [-]^b\, \mathrm{V}_{a-b})\, vu \\ &= (-)^{ab+b}(\mathrm{V}_{a-b} - \mathrm{V}_{a-b+2} + \dots + [-]^b\, \mathrm{V}_{a+b})\, vu \end{aligned}\right\}. \tag{2}$$

The following are easy deductions

$$\left.\begin{aligned} \tfrac{1}{2}[uv + (-)^{ab} vu] &= (\mathrm{V}_{a+b} + \mathrm{V}_{a+b-4} + \mathrm{V}_{a+b-8} + \dots)\, uv \\ \tfrac{1}{2}[uv - (-)^{ab} vu] &= (\mathrm{V}_{a+b-2} + \mathrm{V}_{a+b-6} + \dots)\, uv \end{aligned}\right\}, \tag{3}$$

and these may be written in the reverse order with $(-)^{ab+b}$ in place of $(-)^{ab}$. From (3)

$$\left.\begin{aligned} \tfrac{1}{2}(u\omega + \omega u) &= (\mathrm{V}_{a+2} + \mathrm{V}_{a-2})\, u\omega = (\mathrm{V}_{a+2} + \mathrm{V}_{a-2})\, \omega u \\ \tfrac{1}{2}(u\omega - \omega u) &= \mathrm{V}_a \,.\, u\omega = -\mathrm{V}_a \,.\, \omega u \end{aligned}\right\}, \tag{4}$$

which has important applications to infinitesimal rotations. Other special cases are

$$u^2 = (\mathrm{V}_0+\mathrm{V}_4+\mathrm{V}_8+\ldots)\,u^2, \tag{5}$$

$$\omega^2 = (\mathrm{V}_0+\mathrm{V}_4)\,\omega^2, \tag{6}$$

$$\alpha\beta = (\mathrm{V}_0+\mathrm{V}_2)\,\alpha\beta = (\mathrm{V}_0-\mathrm{V}_2)\,\beta\alpha, \tag{7}$$

$$\mathrm{V}_2\alpha\beta = -\mathrm{V}_2\beta\alpha, \qquad \mathrm{V}_0\alpha\beta = \mathrm{V}_0\beta\alpha. \tag{8}$$

$$\mathrm{V}_2\alpha\beta = \tfrac{1}{2}(\alpha\beta-\beta\alpha), \qquad \mathrm{V}_0\alpha\beta = \tfrac{1}{2}(\alpha\beta+\beta\alpha). \tag{9}$$

§ 4. *Rotations of Three Kinds.*—Suppose we find n multenions $p_1, p_2 \ldots p_n$ such that for all purposes of Laws I to IV they may take the place of $i_1, i_2 \ldots i_n$, are we to call the substitution of the multenions p for the vectors i a rotation? In Euclidean space of three dimensions we should not always so affirm. If i, j, k should represent, as sometimes in quaternions they do, three directions in a rigid body we could not move the body so that i became $-i$ and j and k remained unaltered; but they would continue to satisfy the conditions of the laws. We should call the transformation a perversion or reflection rather than a rotation. As the algebra of multenions is (n even) a full linity algebra, or (n odd) a sub-algebra of such a linity algebra, we will assume that given q, q^{-1} can be obtained uniquely in general so that $qq^{-1} = 1 = q^{-1}q$ (though in certain singular cases we have to say instead that q_0 can be found so that $qq_0 = 0 = q_0q$). In the general case we will say that q is vertible (*i.e.*, it has an inverse q^{-1}); in the singular case that q is non-vertible.

Theorem concerning *multenion* rotation. (1) *When q is a given vertible multenion, then in the four fundamental laws the primitive vectors $i_1, i_2, \ldots i_n$ may be replaced respectively by*

$$p_1 = qi_1q^{-1}, \ldots, \qquad p_n = qi_nq^{-1}, \tag{1}$$

and the effect of the replacement is to change any multenion r to

$$r' = qrq^{-1}. \tag{2}$$

(2) *Conversely if $p_1, p_2 \ldots p_n$ are n multenions such that*

$$-1 = p_1^2 = p_2^2 = \ldots = p_n^2, \tag{3}$$

and every pair is anti-commutative, that is

$$p_1p_2+p_2p_1 = 0, \text{ etc.}, \tag{4}$$

and when n is odd

$$p_1p_2\ldots p_n = i_1i_2\ldots i_n, \tag{5}$$

then there exists a vertible multenion, q, for which (1) *and* (2) *are true; r' now meaning the multenion obtained from the given multenion r by replacing $i_1, i_2, \ldots i_n$ by $p_1, p_2, \ldots p_n$ respectively.*

This kind of rotation is scarcely what we seek. For instance, it does not in general leave V_c unchanged, that is, it is not in general true that

$$qV_c r \,.\, q^{-1} = V_c(qrq^{-1}). \tag{6}$$

Multenion rotation qrq^{-1} is operation by a multenion linity $\phi = q(\,)q^{-1}$ such that $\phi'\phi = 1 = \phi\phi'$, and conversely.

Next, imposing the condition that (6) is to be satisfied, we find that q must be the product of any number of vectors, and that if it is such a product (6) is satisfied. We may next find what is the effect of infinitesimally changing the vectors and attempt to build up the finite continuous group of which that infinitesimal transformation is the element. The infinitesimal rotation is the transformation effected by changing q to $q+\frac{1}{2}(\omega q - q\omega)$, where ω is infinitesimal and the corresponding finite rotation is $e^{\omega}(\,)e^{-\omega}$, where ω is finite ω is a perfectly arbitrary ${}_nV_2$, infinitesimal of course for the first case. When ω is not singular ($\omega^2 = 0$) its general form may be put

$$\omega = a\epsilon_1\epsilon_2 + b\epsilon_3\epsilon_4 + \ldots,$$

where $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 \ldots$ are orthogonal unit vectors

$$e^{\omega} = (\cos a + \epsilon_1\epsilon_2 \sin a)(\cos b + \epsilon_3\epsilon_4 \sin b)\ldots$$

and the factors here are commutative. Putting the last in the form

$$e^{\omega} = \epsilon_1(\epsilon_1^{-1}\cos a + \epsilon_2 \sin a)\,\epsilon_3(\epsilon_3^{-1}\cos b + \epsilon_4 \sin b)\ldots,$$

we see that e^{ω} is the product of an *even* number of vectors. The general form of ω when it is singular is not simple, but I have a complete solution before me in MS.

To show that the infinitesimal rotation is as stated will not occupy too much space. Suppose

$$r' = \alpha\beta\ldots\lambda \,.\, r \,.\, \lambda^{-1}\ldots\beta^{-1}\alpha^{-1} = prp^{-1},$$

where the vectors $\alpha, \beta, \gamma, \ldots$ are assumed to be unit vectors, as that does not affect the transformation. Let $\alpha, \beta, \ldots$ change to $\alpha + d\alpha$, $\beta + d\beta, \ldots$, still remaining unit vectors.

$$dr' = d(prp^{-1}) = dpp^{-1} \,.\, r' - r'dpp^{-1}.$$

That dpp^{-1} is a ${}_nV_2$ is straightforwardly proved thus

$$dpp^{-1} = d\alpha\alpha^{-1} + \alpha \,.\, d\beta\beta^{-1} \,.\, \alpha^{-1} + \alpha\beta \,.\, d\gamma\gamma^{-1} \,.\, \beta^{-1}\alpha^{-1} + \ldots.$$

Every term here is a ${}_nV_2$; for example, consider the third term. Since γ and $\gamma + d\gamma$ are unit vectors, $V_0 d\gamma\gamma^{-1} = 0$, and therefore $d\gamma\gamma^{-1}$ is a ${}_nV_2$. It follows that the third term is a ${}_nV_2$. If we put u, u' for r, r' and $\frac{1}{2}\,d\omega$ for dpp^{-1},

$$du' = \tfrac{1}{2}(d\omega u' - u'd\omega) = V_0 d\omega u'.$$

This is the very simple, very useful form in which true infinitesimal rotation presents itself in our present subject. A rotation of vectors ρ is a linity $\phi\rho$ of ρ such that $\phi'\phi = 1$.

When without qualification we speak of rotation we shall understand it to mean $e^{\omega}(\ )e^{-\omega}$ (a continuous group).

The skew part of any vector linity, ϕ, is of the form of an infinitesimal rotation. Thus, introduce the ζ notation of quaternions,* by which

$$\mathrm{F}(\zeta, \zeta) = \sum_{c=1}^{n} \mathrm{F}(i_c, i_c), \tag{7}$$

where F is bilinear in its two members. We have

$$\rho = -\zeta \mathrm{V}_0 \zeta\rho, \tag{8}$$

$$\phi'\rho = -\zeta \mathrm{V}_0 \zeta\phi'\rho = -\zeta \mathrm{V}_0 \rho\phi\zeta. \tag{9}$$

The skew part is

$$\tfrac{1}{2}(\phi-\phi')\rho = \tfrac{1}{2}\mathrm{V}_1(\mathrm{V}_2\zeta\phi\zeta)\rho = \mathrm{V}_1\omega\rho. \tag{10}$$

Here we have assumed that

$$\mathrm{V}_1(\mathrm{V}_2\alpha\beta)\gamma = \alpha\mathrm{V}_0\beta\gamma - \beta\mathrm{V}_0\gamma\alpha,$$

much as in quaternions. This is not difficult to prove, and we shall in § 7 enunciate much more general theorems of which this is a particular case. The pure part does not similarly reduce. We have

$$\tfrac{1}{2}(\phi+\phi')\rho = -\tfrac{1}{2}(\zeta\mathrm{V}_0\rho\phi\zeta + \phi\zeta\mathrm{V}_0\rho\zeta). \tag{11}$$

§ 5. *Vectoriums.* Let

$$\alpha^{(a)} = \alpha_1\alpha_2\ldots\alpha_a, \quad \beta^{(b)} = \beta_1\beta_2\ldots\beta_b, \quad \gamma^{(c)} = \gamma_1\gamma_2\ldots\gamma_c.$$

Then in the expression $\mathrm{V}_{a+b+c}\ (\alpha^{(a)}\beta^{(b)}\gamma^{(c)}$ it is obvious by reduction to primitive units that within the brackets we may at will write either $\alpha^{(a)}$ or $\mathrm{V}_a\alpha^{(a)}$, and $\beta^{(b)}$ or $\mathrm{V}_b\beta^{(b)}$, and $\gamma^{(c)}$ or $\mathrm{V}_c\gamma^{(c)}$; the value of the expression is unaltered by such insertion or removal of V_a, V_b, or V_c. The statement, of course, is in general only true of this particular part, V_{a+b+c} of the product of highest homogeneity. Next, it is similarly evident that if we change, say, α_2 to $x\alpha_2 + y\alpha_2'$ we simply alter the corresponding vectorium similarly, that is

$$\mathrm{V}_a\alpha_1(x\alpha_2+y\alpha_2')\alpha_3\ldots\alpha_a = x\mathrm{V}_a\alpha_1\alpha_2\alpha_3\ldots\alpha_a + y\mathrm{V}_a\alpha_1\alpha_2'\alpha_3\ldots\alpha_a.$$

Again, if in the vectorium we interchange two consecutive members such as α_2, α_3 we merely change the sign of the whole. For we may successively write instead of $\alpha_2\alpha_3$

$$\alpha_2\alpha_3, \quad \mathrm{V}_2\alpha_2\alpha_3, \quad -\mathrm{V}_2\alpha_3\alpha_2, \quad -\alpha_3\alpha_2.$$

Finally, by a familiar simple process we may interchange any two members by successive interchanges of neighbours, and as a final result, again, the vectorium changes in sign merely. Thus $\mathrm{V}_a\alpha^{(a)}$ is "combinatorial." The common simple properties of a combinatorial multilinear function are probably well known to my readers and may be assumed. [Grassmann apparently first, in a perfectly general manner, appreciated the value of, and fully established these properties. Indeed, $V_a\alpha^{(a)}$ is both algebraically and geometrically identical in meaning with Grassmann's combinatorial "product." I should like for a second time to express my unbounded admiration for Grassmann's great pioneer work, though I cannot help believing that the actual form of the matters he dealt with has been put out of court by later developments.]

---

* See 'Utility of Quaternions.'

We may note the following theorem in passing. It is our form of one of Grassmann's products. If $\acute{\rho}$ is a vectorium of homogeneity, g, then

$$V_{a+b+c}(V_{g+a+b}\acute{\rho}uv)(V_{g+c}w\acute{\rho}) = V_{a+b+c}(V_{g+a}\acute{\rho}u)(V_{g+b+c}vw\acute{\rho}).$$

Of u, v, w the middle one, v, may be passed over from the first to the second bracket-pair.

The following is wanted immediately and is constantly in request. It may be thought of as a passage into or out of the operation of a $V_a(\,)$ [whether as pre-factor or post-factor] of either $\acute{\upsilon}$ or p_n. We have

$$\acute{\upsilon}V_aq = V_{n-a}\acute{\upsilon}q, \qquad V_aq\,.\,\acute{\upsilon} = V_{n-a}(q\acute{\upsilon}).$$

Also, if p_n is any multenion of homogeneity, n, $p_n = x\acute{\upsilon}$, and, therefore, in the present process, p_n behaves exactly as does $\acute{\upsilon}$, that is

$$p_nV_aq = V_{n-a}p_nq, \qquad V_aq\,.\,p_n = V_{n-a}\,.\,qp_n.$$

Suppose there are n given linearly independent vectors $\alpha_1, \alpha_2, \dots \alpha_n$. Let $\alpha^{(a)}$ be the product of any a of them, and $\alpha^{(n-a)}$ the product of the rest in such sequence that $V_n\alpha^{(a)}\alpha^{(n-a)}$ always retains the same value. Let

$$\left.\begin{aligned}\acute{a} &= V_a\alpha^{(a)}\\ \grave{a} &= V_{n-a}\alpha^{(n-a)}\,.\,V_n{}^{-1}\alpha^{(a)}\alpha^{(n-a)}\end{aligned}\right\}. \tag{1}$$

[By a convenient, strictly speaking indefensible, custom, since $V_n{}^{-1}$ really $= \infty$, we mean by $V_n{}^{-1}\alpha^{(a)}\alpha^{(n-a)}$

$$(V_n\alpha^{(a)}\alpha^{(n-a)})^{-1}].$$

There are 2^n linearly independent [proof of linear independence here omitted] values of $\acute{a}$, and they will be called the 2^n vectoriums *formed from* $\alpha_1, \alpha_2, \dots \alpha_n$. [The term may, of course, be used when $\alpha_1, \alpha_2, \dots \alpha_n$ are not linearly independent.] The 2^n corresponding values of $\grave{a}$, which are also linearly independent, will be called the normal reciprocals of the vectoriums, $\acute{a}$. They are so called because

$$V_0\acute{a}\grave{a} = 1, \qquad V_0\acute{a}_1\grave{a} = 0,$$

$\acute{a}_1$ being a second (not corresponding to $\grave{a}$) $\acute{a}$. [Thus, when $n = 4$, there are sixteen equations of the type $V_0\acute{a}\mathring{a} = 1$ and 240 equations of the type $V_0\acute{a}_1\mathring{a} = 0$.] Both equations are proved quite simply from (1) by passing $p_n = V_n{}^{-1}\alpha^{(a)}\alpha^{(n-a)}$ out of the operation of V_0 in $V_0\acute{a}\mathring{a}$.

Let now $q = \Sigma x\acute{a} = \Sigma y\mathring{a}$. From the normal reciprocal relations we at once get $x = V_0q\mathring{a}$, $y = V_0q\acute{a}$. Hence

$$\left.\begin{aligned} q &= \Sigma\acute{a}V_0q\mathring{a} = \Sigma V_a\alpha^{(a)} \,.\, V_nq\alpha^{(n-a)} \,.\, V_n{}^{-1}\alpha^{(a)}\alpha^{(n-a)} \\ &= \Sigma\mathring{a}V_0q\acute{a} = \Sigma V_0(qV_a\alpha^{(a)}) \,.\, V_{n-a}\alpha^{(n-a)} \,.\, V_n{}^{-1}\alpha^{(a)}\alpha^{(n-a)} \end{aligned}\right\}, \tag{2}$$

the last result in the first line coming from a new application of the p_n theorem. If in (2) we put $q = u$, getting a more particular result, we should at once see that (2) is analogous to the quaternion theorems

$$\rho S\alpha\beta\gamma = \alpha S\beta\gamma\rho + \ldots = V\beta\gamma S\alpha\rho + \ldots.$$

A particular case ($a = 1$) of (1) is

$$\left.\begin{aligned} \hat{a}_1 &= V_{n-1}\alpha_2\alpha_3\ldots\alpha_n \,.\, V_n{}^{-1}\alpha_1\alpha_2\ldots\alpha_3 \\ &= -V_{n-1}\alpha_1\alpha_3\ldots\alpha_n \,.\, V_n{}^{-1}\alpha_1\alpha_2\ldots\alpha_3 = \text{etc.} \end{aligned}\right\}. \tag{3}$$

This is clearly the solution of the problem. Given

$$S\alpha_c\hat{a}_c = 1, \qquad c = 1, 2, \ldots n,$$
$$S\alpha_b\hat{a}_c = 0, \qquad b \neq c;\ b, c = 1, 2, \ldots n,$$

what are the vectors $\hat{a}_c$ in terms of the vectors α_c? And, of course, we have a precisely similar solution for α_c in terms of $\hat{a}_c$. It appears, then, that the relations between the quantities $\acute{a}$ and the quantities $\mathring{a}$ are symmetrical. No doubt we have here established this if only it be evident what is the precise significance of the word "symmetrical" in this connection. The need of the caution may be illustrated thus. If we put $\acute{a} = V_a\alpha_1\alpha_2 \ldots \alpha_n$, we are tempted to say that $\mathring{a}$ must be $V_a\hat{a}_1\hat{a}_2 \ldots \hat{a}_a$, but we shall find that we do not thereby always obtain $S\acute{a}\mathring{a} = 1$. The true value of $\mathring{a}$ is

$$\left.\begin{aligned} \mathring{a} &= V_a\hat{a}_a\hat{a}_{a-1}\ldots\hat{a}_1 = QV_a\hat{a}_1\hat{a}_2\ldots\hat{a}_a \\ \text{when}\qquad \acute{a} &= V_a\alpha_1\alpha_2\ldots\alpha_a \end{aligned}\right\}. \tag{4}$$

If the α_c are taken as the i_c, it is obvious that the $\acute{a}$ are the primitive vectoriums, and, from the normal reciprocal relations, it is then obvious that the $\mathring{a}$ are also the primitive vectoriums (with different signs in accordance with)

$$\mathring{a} = K\acute{a}.$$

Return to the ζ pairs of (7) of § 4. Let $F(q, r)$ be bilinear in two multenions. From the definition of ζ, it is easily seen that

$$F(V_a\zeta_1\zeta_2\ldots\zeta_a, V_a\zeta_1\zeta_2\ldots\zeta_a) = a\,!\ \Sigma F(\acute{i}_a, \mathring{i}_a) = (-)^{\frac{1}{2}a(a+1)}a\,!\ \Sigma F(\acute{i}_a, \acute{i}_a{}^{-1}).$$

Abbreviate thus

$$\zeta_1\zeta_2\ldots\zeta_a = \zeta^{(a)}, \quad (-)^{\frac{1}{2}a(a+1)}/a\,! = a'. \tag{5}$$

[Thus $0' = 1$, $1' = -1$, $2' = -1/2!$, $3' = +1/3!$ etc., the signs running $+--++--++--\ldots$.]

Thus $$\Sigma F(i_a, i_a^{-1}) = a'F(V_a\zeta^{(a)}, V_a\zeta^{(a)}). \tag{6}$$

From (6) $$u = \Sigma i_a V_0 u i_a^{-1} = a'V_a\zeta^{(a)} \,.\, V_0 u\zeta^{(a)}. \tag{7}$$

In (6) put the second $V_a\zeta^{(a)} = \Sigma \dot{a} V_0 \mathring{a}\zeta^{(a)}$. Thus,

$$a'F(V_a\zeta^{(a)}, V_a\zeta^{(a)}) = a'\Sigma(V_a\zeta^{(a)} V_0 \mathring{a}\zeta^{(a)}, \dot{a}) = \Sigma F(\mathring{a}, \dot{a}).$$

We might have modified the first $V_a\zeta^{(a)}$ instead of the second. Thus collecting results

$$\left.\begin{aligned} a'F(V_a\zeta^{(a)}, V_a\zeta^{(a)}) &= \Sigma F(i_a, i_a^{-1}) \\ &= \Sigma F(\mathring{a}, \dot{a}) = \Sigma F(\dot{a}, \mathring{a}) \end{aligned}\right\} \tag{8}$$

(${}_nC_a$ terms in each summation.)

Just as in quaternions $F(\zeta, \phi\zeta) = F(\phi'\zeta, \zeta)$ so now if ϕ is a ${}_nV_a$ linity

$$a'F(V_a\zeta^{(a)}, \phi V_a\zeta^{(a)}) = a'F(\phi' V_a\zeta^{(a)}, V_a\zeta^{(a)}). \tag{9}$$

If ϕ is a multenion linity

$$\sum_{a=0}^{n} a'F(V_a\zeta^{(a)}, \phi V_a\zeta^{(a)}) = \sum_{a=0}^{n} a'F(\phi' V_a\zeta^{(a)}, V_a\zeta^{(a)}). \tag{10}$$

For $a = 0$ we must put $a' = 1$, and for each $V_a\zeta^{(a)}$ we must also put 1 [see (7) above].

§ 6. *Extended Vector Linities.*—The theorems of §§ 6, 7 find continuous applications in the multenion treatment of differential invariants. Such treatment may be considered to begin in § 8, and is explicitly introduced in § 9.

Let ϕ be a given vector linity. The extended linity (a multenion linity) which may for most purposes, but not for all, also be denoted by ϕ, is here defined thus. Let $i_a = i_g i_h \ldots i_l$ be any primitive unit of homogeneity, a, a itself being arbitrary. Then

$$\phi V_a i_a = V_a \phi i_g \phi i_h \ldots \phi i_l.$$

Now change the n primitive vectors $i_1, i_2, \ldots i_n$ to any other n vectors by such steps as replacing i_1 by $xi_1 + yi_2$ (or, if i_1, i_2 have already changed to ϵ_1, ϵ_2, the change should be from ϵ_1 to $x\epsilon_1 + y\epsilon_2$). Such a change simply multiplies each side of the defining equation (when i_1 occurs in it) by x; the meaning is therefore unaltered. It follows that $\phi V_a \alpha^{(a)} = V_a(\phi\alpha)^{(a)}$ where a is arbitrary and the vectors α are arbitrary. If, then, (ϕ), (ψ) are the extended linities of ϕ, ψ, we see that $(\phi)(\psi)$ is the extended linity of $\phi\psi$, for

$$(\phi)(\psi) V_a i_g i_h \ldots i_l = (\phi) V_a \psi i_g \psi i_h \ldots \psi i_l = V_a \phi\psi i_g \phi\psi i_h \ldots \phi\psi i_l.$$

From this again it follows that the extended linity of ϕ^{-1} is $(\phi)^{-1}$.

Let us use either (ϕ) or ϕ for the extended linity of ϕ for the present. What we have just proved may be put

$$(\phi)(\psi) = (\phi\psi),$$
$$(\phi)^{-1} = (\phi^{-1}).$$

We are now about to prove a similar proposition concerning ϕ' the conjugate, namely

$$(\phi)' = (\phi').$$

By (7) of § 5 we have

$$\phi u = a'\phi \mathrm{V}_a\zeta^{(a)}\mathrm{V}_0u\zeta^{(a)},$$

or

$$(\phi)\, u = a'\mathrm{V}_a(\phi\zeta)^{(a)}\mathrm{V}_0u\zeta^{(a)}. \tag{1}$$

Let u, u_0 be two arbitrary multenions of homogeneity a. Operate on (1) by $\mathrm{V}_0u_0\,(\ )$

$$\mathrm{V}_0u_0\,(\phi)\, u = a'\mathrm{V}_0u_0\,(\phi\zeta)^{(a)}\mathrm{V}_0u\zeta^{(a)}.$$

Modifying the left by the fundament rule of the meaning of $(\phi)'$, and modify the right by the rule $\mathrm{F}(\zeta, \phi\zeta) = \mathrm{F}(\phi'\zeta, \zeta)$ applicable to a vector linity. We obtain

$$\mathrm{V}_0u\,(\phi)'u_0 = a'\mathrm{V}_0u_0\zeta^{(a)}\mathrm{V}_0u\,(\phi'\zeta)^{(a)}.$$

Since u is arbitrary, by (16) of § 3

$$(\phi)'u_0 = a'\mathrm{V}_a(\phi'\zeta)^{(a)}\mathrm{V}_0u_0\zeta^{(a)},$$

which by comparison with (1) proves the assertion that $(\phi)' = (\phi')$.

The following theorems are also true

$$\phi\mathrm{V}_{a+b}uv = \mathrm{V}_{a+b}\phi u\,\phi v, \tag{2}$$

$$\phi\mathrm{V}_{a-b}uv = \mathrm{V}_{a-b}\phi u\phi'^{-1}v, \tag{3}$$

when ϕ is non-vertible we must change v to $\phi'v$. An alternative, less convenient than (3) generally, is thus

$$\phi\mathrm{V}_{a-b}u\phi'v = \mathrm{V}_{a-b}\phi uv, \tag{4}$$

but this alternative is true without qualification.

$\acute{v}$ having its usual meaning $(i_1, i_2, \ldots i_n)$ it is easy to see that $\phi\acute{v} = m\acute{v}$ where m is the determinant of ϕ. Putting $u = \acute{v}$ in (3) and (4) we obtain

$$\left.\begin{aligned} \acute{v}^{-1}\phi(\acute{v}v) &= m\phi'^{-1}v \\ \text{or}\qquad \acute{v}^{-1}\phi(\acute{v}\phi'v) &= mv \end{aligned}\right\}. \tag{5}$$

[*Added* ***February***, 1921.—Closely connected with the extension in § 6 from vector operation to multenion operation, of a given vector linity, is a second kind of such extension. If ϕ_{E} is the extended (or, as I now prefer to say *extensive*) form of ϕ of § 6, its fundamental property, showing its dependence (in a *multiple* manner) on ϕ, is that

$$\phi_{\mathrm{E}} \mathrm{V}_a \alpha_1 \alpha_2 \ldots \alpha_a = \mathrm{V}_a \phi\alpha_1 \phi\alpha_2 \ldots \phi\alpha_a. \tag{1}$$

Similarly, if ψ_e is the new extensive form of ϕ, its fundamental property showing its dependence (in an *additive* manner) on ϕ is that

$$\psi_e \mathrm{V}_a \alpha_1 \alpha_2 \ldots \alpha_a = \mathrm{V}_a (\psi\alpha_1)\, \alpha_2 \ldots \alpha_a + \mathrm{V}_a \alpha_1 (\psi\alpha_2)\, \alpha_3 \ldots \alpha_a + \ldots. \tag{2}$$

The dependence is called multiple or additive because

$$(\phi\Phi)_{\mathrm{E}} = \phi_{\mathrm{E}} \Phi_{\mathrm{E}}, \qquad (\psi + \Psi)_e = \psi_e + \Psi_e. \tag{3}$$

We may speak briefly of ϕ_{E} as the extensive of ϕ and ψ_e as the subextensive of ψ (using extensive and subextensive both as nouns and adjectives; compare offensive and defensive; compare also Whitehead's translation of "Ausdehnungslehre" as "Calculus of extension").

For both kinds of extensive it is true, when ϕ_{E} and ψ_e operate on a mere vector, that

$$\phi_{\mathrm{E}} \rho = \phi\rho, \qquad \psi_e \rho = \psi\rho, \tag{4}$$

so that quite frequently it is convenient to use ϕ for ϕ_{E} and on other occasions to use ψ for ψ_e. [We should very rarely with a given ϕ want on a single occasion to use both ϕ_{E} and ϕ_e, but if we do we must have some mark to distinguish them, since they have diverse meanings.]

Although the properties of a subextensive linity may be derived directly from its fundamental property defined above, they may instead be derived from those of an extensive by the methods connecting an infinitesimal with a finite transformation. Thus we may put

$$\left.\begin{aligned} \phi &= 1 + \dot{\phi}\,dt = 1 + \psi\,dt \\ (1 + \psi\,dt)_{\mathrm{E}} = \phi_{\mathrm{E}} &= (1 + \dot{\phi}\,dt)_{\mathrm{E}} = 1 + \psi_e\,dt \end{aligned}\right\}, \tag{5}$$

where dt is an arbitrary infinitesimal scalar. Thus from former results we at once have

$$\mathrm{V}_c \psi_e = \psi_e \mathrm{V}_c = \mathrm{V}_c \psi_e \mathrm{V}_c, \tag{6}$$

$$(\psi_e)' = (\psi')_e, \tag{7}$$

$$\psi_e \mathrm{V}_{a+b} uv = \mathrm{V}_{a+b} (\psi_e u \,.\, v + u \psi_e v), \tag{8}$$

$$\psi_e \mathrm{V}_{a-b} uv = \mathrm{V}_{a-b} (\psi_e u \,.\, v - u \psi_e' v), \tag{9}$$

$$\psi_e \dot{v} = \psi_e' \dot{v} = -\dot{v} \,.\, \mathrm{V}_0 \zeta\psi\zeta, \tag{10}$$

$$\psi_e v = -\mathrm{V}_b \psi\zeta \mathrm{V}_{b-1} \zeta v = -\mathrm{V}_b (\mathrm{V}_0 \psi\zeta\zeta \,.\, v - \zeta \mathrm{V}_{b+1} \psi\zeta v), \tag{11}$$

$$\dot{v}^{-1} \psi_e (\dot{v} v) = -(\psi_e' + \mathrm{V}_0 \zeta\psi\zeta)\, v = -\mathrm{V}_b \psi\zeta \mathrm{V}_{b+1} \zeta v, \tag{12}$$

the last being obtained by putting $u = \dot{v}$ in (9) above.

If ϕ and ψ are given vector linities then $\phi_{\mathrm{E}}\psi_{s}\phi_{\mathrm{E}}^{-1}$ is a subextensive that is it is the subextensive of $\phi\psi\phi^{-1}$. For by (11)

$$(\phi\psi\phi^{-1})_s v = -\mathrm{V}_b\phi\psi\phi^{-1}\zeta\mathrm{V}_{b-1}\zeta v = -\phi_{\mathrm{E}}\mathrm{V}_b\psi\zeta\mathrm{V}_{b-1}\zeta\phi_{\mathrm{E}}^{-1}v,$$

by applications of both the equations [see (2) and (3) of § 6],

$$\phi_{\mathrm{E}}\mathrm{V}_{a+c}uw = \mathrm{V}_{a+c}\phi_{\mathrm{E}}u\phi_{\mathrm{E}}w, \qquad \phi_{\mathrm{E}}\mathrm{V}_{c-a}uw = \mathrm{V}_{c-a}\phi_{\mathrm{E}}'^{-1}u\phi_{\mathrm{E}}w.$$

If ϕ is of assigned dimensions, ϕ_s is of the same, but not ϕ_{E}. $\phi_{\mathrm{E}}u$, in so far as its dimensions depend on ϕ_{E} and not on u, are a times the dimensions of ϕ.]

§ 7. *Formulæ Analogous to the Quaternion Formulæ for* $\mathrm{V}\alpha\mathrm{V}\beta\gamma$ *and* $\mathrm{V}\alpha\beta\gamma$.—Let $\alpha_1, \alpha_2, \ldots, \alpha_{a+b}$ be $a+b$ given vectors and let $\alpha^{(a)}$ be the product of any a of them and $\alpha^{(b)}$ the product of the rest in such a sequence that $\mathrm{V}_{a+b}\alpha^{(a)}\alpha^{(b)}$ always has the same value. Then u being as usual

$$\mathrm{V}_b \,.\, u\mathrm{V}_{a+b}\alpha^{(a)}\alpha^{(b)} = \mathrm{V}_b(\mathrm{V}_{a+b}\alpha^{(b)}\alpha^{(a)} \,.\, u) = \Sigma\mathrm{V}_b\alpha^{(b)}\mathrm{V}_0\alpha^{(a)}u. \tag{1}$$

Note that $\alpha^{(a)}$ occupies the middle position throughout. If in this we put $b = n-a$ it easily leads to the last form of (2) of § 5.

u and v being as usual and ρ any vector,

$$\mathrm{V}_{1+(a\sim b)} \,.\, u\mathrm{V}_{b+1}\rho v = \mathrm{V}_{1+(a\sim b)}[\mathrm{V}_{a-1}u\rho \,.\, v + (-)^a\rho\mathrm{V}_{a\sim b}uv], \tag{2}$$

$$\mathrm{V}_{1+(a\sim b)}(\mathrm{V}_{a+1}u\rho \,.\, v) = \mathrm{V}_{1+(a\sim b)}[u\mathrm{V}_{b-1}\rho v + (-)^a\rho\mathrm{V}_{a\sim b}uv], \tag{3}$$

$$\left.\begin{aligned}\mathrm{V}_{1+(a\sim b)}u\rho v &= \mathrm{V}_{1+(a\sim b)}[\mathrm{V}_{a-1}u\rho \,.\, v + u\mathrm{V}_{b-1}\rho v + (-)^a\rho\mathrm{V}_{a\sim b}uv]\\ &= \mathrm{V}_{1+(a\sim b)}[\mathrm{V}_{a+1}u\rho \,.\, v + u\mathrm{V}_{b+1}\rho v - (-)^a\rho\mathrm{V}_{a\sim b}uv]\end{aligned}\right\}, \tag{4}$$

$$\left.\begin{aligned}\mathrm{V}_{a+b-1}u\rho v &= \mathrm{V}_{a+b-1}[\mathrm{V}_{a-1}u\rho \,.\, v + u\mathrm{V}_{b+1}\rho v + (-)^a\rho\mathrm{V}_{a+b}uv]\\ &= \mathrm{V}_{a+b-1}[\mathrm{V}_{a+1}u\rho \,.\, v + u\mathrm{V}_{b-1}\rho v - (-)^a\rho\mathrm{V}_{a+b}uv]\end{aligned}\right\}. \tag{5}$$

To make practical applications of multenions it is absolutely essential that facility should be gained in using these formulæ. The following comments are to aid in the acquirement of the requisite facility. I do not attempt to work without a list of at least the four equations (4) and (5) before me. Each of those four generates two others. (2) and (3) are those two derived from the first equation (4). If the second term on the right of this equation (call it (4*a*)) is removed to the left it can be made to merge with the left and equation (2) is produced. Similarly, if the first term on the right of (4*a*) is removed to the left equation (3) is produced. Now as after a little practice, with (4*a*) before the eye, the whole of this process can be gone through mentally and the desired modification of an expression under treatment can forthwith be written down, (2) and (3) are strictly speaking not required in the working list. Each of the four equations (4) and (5) is similar in the points described to (4*a*). The reader who desires to acquire this facility could not do better than begin by writing out all the twelve equations implied.

Attend now only to (4) and (5). If the two equations (4) are added together one reaches a mere identity almost immediately. Therefore, either of the equations (4) readily follows from the other. Exactly similarly either of the equations (5) follows from the other. To obtain one of (4) from one of (5), place $\acute{v}$ as a post factor to one of (5) and then pass it into the various operators $Vx(\ )$ till it reaches v, noting how those operators must be changed in the process. When $\acute{v}$ has come to juxtaposition with v, change $v\acute{v}$ into w so that $n-b=c$ and get rid of b, leaving a and c. The completion of the verification is simple. I have verified that it is possible to prove either of (5) by reduction to primitive units. I have found no simple proof. Actually I established one of the twelve equations by taking ρ to be one of the vectors α_c, say α_{a+b} in (1). Perhaps this is the simplest method.

Because the space I have assigned myself does not permit otherwise, I must now omit several chapters in the MS. treatise from which I have been summarising, but I find one matter requires insertion before I pass to the part of the subject lending itself to applications to differential invariants.

If ϕ is a multenion linity its characteristic function $chf_\phi(x)$ is a certain n-ic in the arbitrary scalar symbol x. Another name for it is that it is the determinant of $x-\phi$, and from the point of view suggested it may be denoted by $|x-\phi|$. In our present notation this is given by the identity

$$\left.\begin{aligned} chf_\phi(x) &\equiv |x-\phi| \\ &\equiv n'\mathrm{V}_n\zeta^{(n)}\,.\,\mathrm{V}_n[(x-\phi)\,\zeta]^{(n)} \\ &\equiv x^n - m_1 x^{n-1} + \ldots + (-)^n m_n \end{aligned}\right\}, \tag{6}$$

[n' is the same function of n as a' is of a in (5) of § 5, that is $n' = (-)^{\frac{1}{2}n(n+1)}/n!$ m_n here and in quaternions is generally written as m]. Several forms derived from (6) are useful in the present subject. They are expressions for m_a of which it is to be remembered that m_n is a specially important case. There is (1) a fundamental form in ζ, (2) a form involving n given linearly independent vectors $\alpha_1, \alpha_2, \ldots \alpha_n$, (3) a form dependent on the general expression for ϕ

$$\phi\epsilon = -\beta_1\mathrm{V}_0\epsilon\alpha_1 - \beta_2\mathrm{V}_0\epsilon\alpha_2 - \ldots = -\Sigma\beta\mathrm{V}_0\epsilon\alpha, \tag{7}$$

and (4) the particular case of the last for which

$$\phi\epsilon = -\mathrm{V}_0\epsilon\nabla. \tag{8}$$

In the order mentioned, these forms are

$$\left.\begin{aligned} m_a &= a'\mathrm{V}_0(\mathrm{V}_a\zeta^{(a)}\phi\mathrm{V}_a\zeta^{(a)}) \\ &= \Sigma\,.\,\mathrm{V}_n[\mathrm{V}_a(\phi\alpha)^{(a)}\,.\,\mathrm{V}_{n-a}\alpha^{(n-a)}]\,.\,\mathrm{V}_n{}^{-1}\alpha^{(a)}\alpha^{(n-a)} \\ &= \Sigma\,.\,\mathrm{V}_0\mathrm{V}_a\alpha^{(a)}\mathrm{K}\mathrm{V}_a\beta^{(a)} \\ &= a'\mathrm{V}_0(\mathrm{V}_a\nabla^{(a)}\mathrm{V}_a\epsilon^{(a)}) \end{aligned}\right\}. \tag{9}$$

The last three are all deducible from the first. Another form could at once be written down from the first, involving the vectors α and the vectors $\hat{a}$.

For our applications below the fourth form is important. For this case

$$\mathrm{F}(\zeta, \phi\zeta) = \mathrm{F}(\nabla_1, \epsilon_1). \tag{10}$$

From this
$$\mathrm{V}_0\zeta\phi\zeta = \mathrm{V}_0\nabla\epsilon, \qquad \mathrm{V}_2\zeta\phi\zeta = \mathrm{V}_2\nabla\epsilon, \tag{11}$$

$$|\phi| = m_n = n'\mathrm{V}_0 \,.\, \mathrm{V}_n\nabla^{(n)}\mathrm{V}_n\epsilon^{(n)}. \tag{12}$$

[*Added February*, 1921.—The formulæ analogous to the quaternion formula for $\mathrm{V}\alpha\mathrm{V}\beta\gamma$ are here greatly generalised and reduced to a form in which they are easy to remember and to apply in particular cases.

In accordance with (3) of § 3 let

$$\left.\begin{aligned} \tfrac{1}{2}[uv+(-)^{ab}vu] &= (\mathrm{V}_{a+b}+\mathrm{V}_{a+b-4}+\mathrm{V}_{a+b-8}+\ldots)\,uv = \Lambda uv \\ \tfrac{1}{2}[uv-(-)^{ab}vu] &= (\mathrm{V}_{a+b-2}+\mathrm{V}_{a+b-6}+\ldots)\,uv = \Lambda' uv \end{aligned}\right\}. \tag{1}$$

The parts of Λuv, and Λuv itself, may be said to be concordant with V_{a+b}, those of $\Lambda' uv$ concordant with V_{a+b-2}. Similarly, let the parts of uvw, as follows,

$$\left.\begin{aligned} \tfrac{1}{2}(uvw+(-)^{bc+ca+ab}wvu) &= \Lambda uvw \\ \tfrac{1}{2}(uvw-(-)^{bc+ca+ab}wvu) &= \Lambda' uvw \end{aligned}\right\}, \tag{2}$$

be called concordant with V_{a+b+c} and $\mathrm{V}_{a+b+c-2}$ respectively. [That Λuvw as defined by (2) is the same as $(\mathrm{V}_{a+b+c}+\mathrm{V}_{a+b+c-4}+\ldots)\,uvw$ may be proved from the properties of K.

$$\mathrm{K}(uvw) = \mathrm{K}w\mathrm{K}v\mathrm{K}u = (-)^{\frac{1}{2}\Sigma a(a+1)}wvu.$$

Hence
$$2\Lambda(uvw) = \{1+(-)^{\frac{1}{2}(a+b+c)(a+b+c+1)}\mathrm{K}\}\,uvw].$$

It is important for present purposes to note the various forms of such an expression as

$$(\Lambda \text{ or } \Lambda') \,.\, u(\Lambda \text{ or } \Lambda')\,vw,$$

obtained by mere commutations, and especially what are the commutation factors in the various cases. There are two standard cases from which all others may be readily derived. They are

$$\Lambda vu = (-)^{ab}\Lambda uv, \qquad \Lambda(\Lambda vw \,.\, u) = (-)^{a(b+c)}\Lambda(u\Lambda vw),$$

and we express them by saying that $(-)^{ab}$ is the commutation factor of Λuv, and that $(-)^{a(b+c)}$ is the commutation factor for u and Λvw in $\Lambda(u\Lambda vw)$. The addition of an accent to any Λ clearly multiplies the commutation factor by -1. Thus

$$\begin{aligned} \Lambda' vu &= -(-)^{ab}\Lambda' uv, \\ \Lambda(\Lambda' vw \,.\, u) &= -(-)^{a(b+c)}\Lambda(u\Lambda' vw), \\ \Lambda'(\Lambda' vw \,.\, u) &= (-)^{a(b+c)}\Lambda'(u\Lambda' vw), \text{ etc.} \end{aligned}$$

Thus, if x, y, are independently equal either to 0 or 1, and if we put $\Lambda^{(x)}$ for Λ or Λ' according as x is 0 or 1, and similarly for $\Lambda^{(y)}$, we have by commutations four forms of $\Lambda^{(x)} . u\Lambda^{(y)}vw$, namely,

$$\begin{aligned}\Lambda^{(x)}(u\Lambda^{(y)}vw) &= (-)^{y+bc}\Lambda^{(x)}(u\Lambda^{(y)}wv)\\ &= (-)^{x+y+a(b+c)}\Lambda^{(x)}(\Lambda^{(y)}vw\,.\,u) = (-)^{x+bc+ca+ab}\Lambda^{(x)}(\Lambda^{(y)}wv\,.\,u). \qquad (3)\end{aligned}$$

It is clear that for all present purposes u, v, w may be generalized in meaning. Hitherto they have meant multenions of the forms V_ap, V_bq, V_cr; but they may here be extended to mean multenions concordant with V_a, V_b, V_c, that is

$$\left.\begin{aligned} u &= (V_a+V_{a+4}+V_{a-4}+V_{a+8}+V_{a-8}+\ldots)p,\\ v &= (V_b+V_{b+4}+V_{b-4}+\ldots)q, \quad w = (V_c+V_{c+4}+\ldots)r\end{aligned}\right\}. \qquad (4)$$

Let now Vq stand for any V_xq, a *single* selected part, which is concordant with Λq; and $V'q$ for any single part which is concordant with $\Lambda' q$. Then the following 8 statements are identities, and each of the 8 has 64 variants obtained by such commutations as we have just been considering. Thus, given u, v, w, there are in all 512 variants, and in each of these V, or V′, has in general a multitude of values. In each of the 8 identities we may add together the individual parts indicated by the V or V′, which means that V or V′ may in every case be replaced by Λ or Λ' respectively. For standard reference forms those written seem preferable. The identities are in meaning (not separately, but taken all together), though not in form, symmetrical in u, v, w. This will appear in the course of the proof given later.

$$\left.\begin{aligned} V(\Lambda uv\,.\,w) &= V(u\Lambda vw+(-)^{ab}v\Lambda' uw)\\ &= V(u\Lambda' vw+(-)^{ab}v\Lambda uw),\\ V(\Lambda' uv\,.\,w) &= V(u\Lambda vw-(-)^{ab}v\Lambda uw)\\ &= V(u\Lambda' vw-(-)^{ab}v\Lambda' uw),\\ V'(\Lambda uv\,.\,w) &= V'(u\Lambda vw+(-)^{ab}v\Lambda uw)\\ &= V'(u\Lambda' vw+(-)^{ab}v\Lambda' uw),\\ V'(\Lambda' uv\,.\,w) &= V'(u\Lambda vw-(-)^{ab}v\Lambda' uw)\\ &= V'(u\Lambda' vw-(-)^{ab}v\Lambda uw)\end{aligned}\right\}. \qquad (5)$$

The following remarks will probably render these easy to remember and apply: (1) Counting the accent of a V′ only *once* when it occurs in an equation, the number of accents in each of the eight equations of three terms is odd, that is, it is 1 or 3. (2) In these standard forms the sequence u, v, w is adopted as far as practicable, that is to say, wholly in the term on the left and the first term on the right; whereas in the second term on the right

v is forced away from the middle position, the sequence u, w is still maintained. (3) When these rules of sequence are adopted the signs to be given to each term are specially easy of determination; when terms on opposite sides of the equation have the same sequence, or by commutation are brought to the same sequence, they have or are brought to have the same sign; and when on the same side to have opposite signs. Thus, as written in (5), the term on the left and the first term on the right, having the same sequence, have the same (plus) sign; and again, the term on the left and the second term on the right are brought to the same sequence by commuting u and v on the left. This last determines the sign, $\pm(-)^{ab}$ in the second term on the right.

We will now throw the eight separate cases of (5) into one form, and henceforth replace V and V′ by Λ and Λ′. Let g, x, y, z each independently be equal to 0 or 1. Then all the cases of (5) are included in

$$\left.\begin{aligned} &\Lambda^{(g)}(\Lambda^{(z)}uv\,.\,w) = \Lambda^{(g)}(u\Lambda^{(x)}vw + (-)^{z+ab}v\Lambda^{(y)}uw) \\ \text{provided}\qquad &g+x+y+z = 1 \text{ or } 3 \end{aligned}\right\}. \tag{6}$$

Put x', y', z' for the commutation factors of $\Lambda^{(x)}vw$, $\Lambda^{(y)}wu$, $\Lambda^{(z)}uv$ respectively, that is

$$x' = (-)^{x+bc}, \qquad y' = (-)^{y+ca}, \qquad z' = (-)^{z+ab}. \tag{7}$$

It is easy from (3) to prove that $-y'z'$, $-z'x'$, $-x'y'$ are the commutation factors for u and $\Lambda^{(x)}vw$ in $\Lambda^{(g)}.\,u\Lambda^{(x)}vw$; and for the like single symbol and coupled pair in $\Lambda^{(g)}.\,v\Lambda^{(y)}wu$ and $\Lambda^{(g)}.\,w\Lambda^{(z)}uv$ respectively.

To see that (6) is symmetrical in u, v, w, make the commutations necessary to obtain the cyclic order u, v, w throughout; first with the single symbol in the first place throughout, and secondly in the last place throughout. Also put all the terms on one side of the equation and write the three terms in the order for which in the first term u occupies the middle position, in the second v does so, and in the third w. Notice that the coefficients x', y', z' thus come in both arrangements to have their alphabetic order. Thus

$$\Lambda^{(g)}(x'w\Lambda^{(z)}uv + y'u\Lambda^{(x)}vw + z'v\Lambda^{(y)}wu) = 0, \tag{8}$$

$$\Lambda^{(g)}(x'\Lambda^{(y)}wu\,.\,v + y'\Lambda^{(z)}uv\,.\,w + z'\Lambda^{(x)}vw\,.\,u) = 0. \tag{9}$$

To show that these are equivalent to (6), we have only to verify for any two terms in each equation that when brought to the same sequence they have opposite signs, and this is quite easy to do. [Commute w with $\Lambda^{(z)}uv$ in the first term of (8); it comes to agreement in sequence and disagreement in sign with the second term.]

We will now prove (8). In the first term in (8), $-x'y'$ is one commutation factor and z' is the other. Hence four times this term is

$$\begin{aligned} 4x'\Lambda^{(y)}w\Lambda^{(z)}uv = 2x'(w\Lambda^{(z)}uv - x'y'\Lambda^{(z)}uv\,.\,w) \\ = 2(x'w\Lambda^{(z)}uv - y'\Lambda^{(z)}uv\,.\,w) \\ = x'\,.\,wuv + z'x'\,.\,wvu - y'uvw - y'z'\,.\,vuw. \end{aligned}$$

We see by cyclic change from this that four times the second and third terms are

$$y'\,.\,uvw + x'y'\,.\,uwv - z'\,.\,vwu - z'x'\,.\,wvu$$

and

$$z'\,.\,vwu + y'z'\,.\,vuw - x'\,.\,wuv - x'y'\,.\,uwv.$$

The sum of these three expressions is identically zero, as is seen by inspection.

We can reproduce the equations (5) in other notations. Thus, let

$$\begin{aligned} vu &= Cuv, & wvu &= Cuvw, \\ \tfrac{1}{2}(1+C)(\ ) &= \overline{C}(\ ), & \tfrac{1}{2}(1-C)(\ ) &= \overline{C}'(\ ), \\ \tfrac{1}{2}(1+K)q &= \overline{K}q, & \tfrac{1}{2}(1-K)q &= \overline{K}'q, \\ \tfrac{1}{2}(1+Q)q &= \overline{Q}q, & \tfrac{1}{2}(1-Q)q &= \overline{Q}'q. \end{aligned}$$

Then, similar to (6), the following is true

$$\left.\begin{aligned} \overline{C}^{(g)}(\overline{C}^{(z)}uv\,.\,w) &= \overline{C}^{(g)}(u\overline{C}^{(x)}vw + (-)^{(z)}v\overline{C}^{(y)}uw) \\ \text{provided}\qquad g+x+y+z &= 1 \text{ or } 3 \end{aligned}\right\}. \qquad (10)$$

The application of the rule of signs is rather improved, but in most cases the reverse is decidedly the case in the application of the proviso. In (10) *within the brackets* each C may be replaced by K, or, instead, by Q, but outside the brackets $\overline{C}^{(g)}$ must be retained; also with K or Q in place of C the index z of $(-)^z$ must be changed to

$$z + \tfrac{1}{2}[a(a\pm 1) + b(b\pm 1)],$$

the upper sign being taken with K and the lower with Q.

(10) is more closely analogous than (6) to the quaternion formula for $V\alpha V\beta\gamma$, but we must remember that in quaternions

$$\begin{aligned} \alpha\beta+\beta\alpha &= 2S\alpha\beta, & \alpha\beta\gamma+\gamma\beta\alpha &= 2V\alpha\beta\gamma, \\ \alpha\beta-\beta\alpha &= 2V\alpha\beta, & \alpha\beta\gamma-\gamma\beta\alpha &= 2S\alpha\beta\gamma, \end{aligned}$$

so that inside the brackets $\overline{C}$, $\overline{C}'$ are analogous to S, V, whereas outside they are analogous to V, S.

Of these different forms (5) is without doubt the best for general use, but the others are occasionally convenient.

Putting $\qquad \Lambda - \Lambda' = L \qquad$ and $\qquad Lu = u = Cu,$

we have that C, K, Q, are all strictly retroscriptive that is

$$\mathrm{C}uvw = \mathrm{C}w\mathrm{C}v\mathrm{C}u, \qquad \mathrm{K}pqr = \mathrm{K}r\mathrm{K}q\mathrm{K}p, \text{ etc.},$$

but L is not, though it has the analogous property

$$\mathrm{L}uv = (-)^{ab}\mathrm{L}v\mathrm{L}u, \qquad \mathrm{L}uvw = (-)^{bc+ca+ab}\mathrm{L}w\mathrm{L}v\mathrm{L}u.$$

These last properties are easily extended to any number of symbols, such as u, v, w. To pass to higher numbers than three with (5), or say (6), is much more troublesome. Here is an example of passing to four symbols. Let u' v' be multenions concordant with $\mathrm{V}_{a'}$, $\mathrm{V}_{b'}$. In $\Lambda^{(g)}(\Lambda^{(h)}uv\Lambda^{(h')}v'u')$, [where outside the brackets $\Lambda^{(g)}$ is concordant with $\mathrm{V}_{a+b+a'+b'-2y}$] first treat $\Lambda^{(h')}v'u'$ as a single symbol, and secondly treat $\Lambda^{(h)}uv$ in that manner. We obtain

$$\left.\begin{array}{l} \Lambda^{(g)}(\Lambda^{(h)}uv\Lambda^{(h')}v'u' \\ \qquad = \Lambda^{(g)}(u\Lambda^{(z)} \,.\, z\Lambda^{(h')}v'u' + (-)^{h+ab}v\Lambda^{(y)} \,.\, u\Lambda^{(h')}v'u' \\ \qquad = \Lambda^{(g)}(\Lambda^{(x')} \,.\, [\Lambda^{(h)}uv \,.\, v']u' + (-)^{h'+a'b'}\Lambda^{(y)} \,.\, [\Lambda^{(h)}uv \,.\, u']v') \\ \text{provided both} \qquad g+(h+h')+(z+y) \\ \text{and} \qquad g+(h+h')+(x'+y') \text{ are odd} \end{array}\right\}. \quad (11)$$

In the equation of four terms provided by (11) there would be thirty-two cases corresponding to the eight of (5), because double value sare assignable independently to g, h, h' x, x'. In each of the equations of three terms provided by (11) there would be sixteen cases.]

§ 8.—*Integration.*—Let $x_1i_1 + x_2i_2 + \ldots + x_ni_n = \rho$ stand for an independent variable position vector of a point in Euclidian space of n dimensions. Let a point trace a curve of which an element is $d\rho\,(= d\rho_1)$. Let this curve move and trace a two-dimensional tract. In the process let an element of the path of the element $d\rho\,(= d\rho_1)$ be $d\rho_2$. The element of the two-dimensional region traced hereby is given by $\mathrm{V}_2d\rho_1d\rho_2$. Let this be continued till a $(b+1)$-dimensional tract has been traced.

Let us consider a space integral over the complete boundary (a b-tract) of the $(b+1)$-tract and express it as a space integral over the $(b+1)$-tract. To do this, starting from any point of the $(b+1)$-tract, let an elementary closed b-tract expand till it has traversed the whole $(b+1)$-tract by reaching its boundary. The process could be described in much more detail. We suppose the $(b+1)$-tract to be eventually filled with parallelepipedal elements. First, for each such element the connection between integral over boundary (of element) is expressed as a multiple of $\mathrm{V}_{b+1}d\rho_1d\rho_2 \ldots d\rho_{b+1}$, denoting that element of $(b+1)$-tract. These are summed and in this sum contributions from the common boundary of two elements cancel. The direction indicated by $d\rho_{b+1}$ is to be *into* the region bounded by the element

$$d\dot{\rho}_b = \mathrm{V}_b d\rho_1 d\rho_2 \ldots d\rho_b.$$

With these conventions the following is true

$$\int^{(b)}\!\!\int \phi d\dot{\rho}_b = \int^{(b+1)}\!\!\int \phi_g \mathrm{V}_b d\dot{\rho}_{b+1} \nabla_g, \tag{1}$$

where

$$\nabla = \sum_{c=1}^{n} i_c \mathrm{D}_{x_c}, \tag{2}$$

$\phi d\dot{\rho}_b$ is the element of boundary-integral due to the element of boundary $d\dot{\rho}_b$, and $d\dot{\rho}_{b+1}$ is an element of the $(b+1)$-tract. The form of ϕ is a function of position. The suffix g means that ∇ acts on this function ϕ.

We will now put the integral (1) in a second form. Let $\dot{\nu} = \mathrm{U}d\dot{\rho}_n$ ($\mathrm{U}d\dot{\rho}_n \mathrm{T}d\dot{\rho}_n = d\dot{\rho}_n$, $(\mathrm{T}d\dot{\rho}_n)^2 = \pm (d\dot{\rho}_n)^2$, $\mathrm{T}d\dot{\rho}_n$ a positive real scalar). Also let $d\dot{\rho}_b = \dot{\nu} d\varsigma_{n-b}$. $d\varsigma_0$ is thus a positive scalar, which measures an element of hyper *bulk*. $d\varsigma_1$ is a vector which measures an element of hyper *area*. Write db for $d\varsigma_0$ and $d\alpha$ for $d\varsigma_1$; then

$$\left.\begin{aligned} db &= \mathrm{V}_0 d\alpha\, d\rho_n \\ d\dot{\rho}_n &= \mathrm{V}_n d\dot{\rho}_{n+1} d\rho_n \end{aligned}\right\}. \tag{3}$$

follows from

Since db is positive and $d\rho_n$ points inwards, $d\alpha$ points outwards in harmony with usual convention.

Equation (1) in the $d\varsigma$ notation becomes

$$\int^{(n-a)}\!\!\int \psi d\varsigma_{n-a} = \int^{(n-a+1)}\!\!\int \psi_g \mathrm{V}_a d\varsigma_{a-1} \nabla_g. \tag{4}$$

The extreme cases of (1) and (4) are when $b = 1$ and $a = 1$ that is

$$\int \phi d\rho = \iint \phi_g \mathrm{V}_1 d\dot{\rho}_2 \nabla_g, \tag{5}$$

$$\int^{(n-1)}\!\!\int \psi d\alpha = \int^{(n)}\!\!\int \psi_g \nabla_g. \tag{6}$$

In (1) put $\phi d\dot{\rho}_b =$ a scalar $= \mathrm{V}_0 v d\dot{\rho}_b$ and make similar changes in the other three. Thus (1) and (4) become

$$\int^{(b)}\!\!\int \mathrm{V}_0 v d\dot{\rho}_b = \int^{(b+1)}\!\!\int \mathrm{V}_0 d\dot{\rho}_{b+1} \mathrm{V}_{b+1} \nabla v, \tag{7}$$

$$\int^{(n-a)}\!\!\int \mathrm{V}_0 uk ds_a = \int^{(n-a+1)}\!\!\int \mathrm{V}_0 d\varsigma_{a-1} \mathrm{V}_{a-1} \nabla (ku), \tag{8}$$

uk has been written in place of a mere u for subsequent convenience. k is supposed to be a given scalar function of position of the nature of density. Thus, if by a general strain ρ becomes ρ', $d\rho$ becomes $d\rho'$, and db becomes db', then k becomes k' where

$$kdb = k'db', \tag{9}$$

k is used in the transformation of (6). (5) and (6) become

$$\int \mathrm{V}_0 \sigma d\rho = \iint \mathrm{V}_0 d\dot{\rho}_2 \mathrm{V}_2 \nabla \sigma, \tag{10}$$

$$\int^{(n-1)}\!\!\int \mathrm{V}_0 \tau k d\alpha = \int^{(n)}\!\!\int \mathrm{V}_0 \nabla (k\tau)\, db. \tag{11}$$

We are about to apply these theorems to differential invariants. It may surprise some readers that there is so much concerning invariants which is independent of any conception of differential quadratic forms. I was acquainted with all the theorems here given to the end of § 10 years before I had heard of quadratic differential forms. My first acquaintance with such forms was after the MS. treatise from which I extract them had been completed. That first (and up to this date my only) acquaintance with such forms was through Prof. Wright's delightful Cambridge tract on the subject. How then did they originate? From a study of Maxwell's suggestions concerning intensities and fluxes. Till the last three months I never looked at them from any other point of view. I gave elaborate applications of their three-dimensional form, twenty-seven years ago, to electrical problems.*

§ 9. *Covariants and Contravariants.*—Let ρ' be a vector function of the independent vector ρ. We may think of this as representing a displacement of every point of a medium filling our n-dimensional Euclidean space from the position ρ to the position ρ', or we may think of it in its pure mathematical aspect as a mere mathematical transformation of the coordinates $x_c = -\mathrm{V}_0 i_c\rho$ into the coordinates $x'_c = -\mathrm{V}_0 i_c \rho'$.

Corresponding to an arbitrary increment $d\rho$ of ρ, there is an increment $d\rho'$ of ρ' given by

$$d\rho' = -\mathrm{V}_0 d\rho\nabla \,.\, \rho' = \chi d\rho. \tag{1}$$

We will use χ to denote the linity extended from the vector linity χ which is given by (1). Invariance has to do with relations which remain unchanged in form, in some defined sense or senses, when ρ' is taken as the independent variable instead of ρ. Suppose h is a scalar function of ρ. Then it is also a function of ρ'. Then

$$dh = -\mathrm{V}_0 d\rho\nabla h = -\mathrm{V}_0 d\rho'\nabla' h, \tag{2}$$

where ∇' is $\sum_{c=1}^{n} i_c \mathrm{D}_{x'_c}$. This is a relation of the kind just mentioned and we want to systematise the treatment of such relations. Anybody who reads my paper of 1892 will see how it came about that it should seem a very natural course to use the theorems of integration of § 8 above for the purpose.

We scarcely need to analyse the details of the strain after what has already been said. The ordinary notions of curl, $\mathrm{V}_2\nabla\sigma$, of an intensity, that is a covariant vector, σ, are given by (10) of § 8. The ordinary notions of convergence of a vector flux, $k\tau$, are given by (11) of § 8. [$k\tau$ is a vector flux; τ is a contravariant vector.] The ordinary notions of the half curl, $\frac{1}{2}\mathrm{V}_2\nabla\sigma$, representing the rate of rotation of a fluid whose velocity is σ, are just the same in n dimensions as in three [see (10) of § 4 and what has been said above about infinitesimal rotations]. But more:—it cannot surprise anybody who has seen fluxes and intensities used as a mathematical method that there is no change in the same uses of the same expressions (curl, convergence, etc.) when we pass to an n-manifold of the most general kind. $V_2\nabla\sigma$ and $V_0\nabla(k\tau)$ will have just the same significance in the manifold as in the original Euclidean space; but this is anticipating.

---

* 'Phil. Trans.,' A, 1892, pp. 685–780.

Synopsis of Notation.

	Covariant.	Contravariant.
(1) Vector	σ	τ
(2) Hypervector	v	u
(3) Multenion	s	r
(4) Multenion linity	ξ	ξ^{-1}

We have given an instance of an invariantive statement in (2). Next, consider the integrals (10) and (7) of § 8, which we here reproduce,

$$\left.\begin{aligned} \int V_0\sigma d\rho &= \iint V_0 d\dot{\rho}_2 V_2\nabla\sigma \\ \int^{(b)}\int V_0 v d\dot{\rho}_b &= \int^{(b+1)}\int V_0 d\dot{\rho}_{b+1} V_{b+1}\nabla v \end{aligned}\right\}. \tag{3}$$

When $\rho^{\backprime}$ is the independent variable, precisely these same integrals are to be expressed in the same form, thus:—

$$\left.\begin{aligned} \int V_0\sigma^{\backprime} d\rho^{\backprime} &= \iint V_0 d\dot{\rho}_2^{\backprime} V_2\Delta^{\backprime}\sigma^{\backprime} \\ \int^{(b)}\int V_0 v^{\backprime} d\dot{\rho}_b^{\backprime} &= \int^{(b+1)}\int V_0 d\rho_{b+1}^{\backprime} V_{b+1}\Delta^{\backprime} v^{\backprime} \end{aligned}\right\}. \tag{4}$$

Now, of course, we are compelled to identify our present $d\rho$ and $d\rho^{\backprime}$ with the $d\rho$ and $d\rho^{\backprime}$ of (1), so that $d\rho^{\backprime} = \chi d\rho$. Again, $d\dot{\rho}_b$ meant $V_b d\rho_1 d\rho_2 \ldots d\rho_b$ where each $d\rho_c$ again must be identified with the $d\rho$ of (1), and therefore the corresponding $d\rho^{\backprime}$ with the $d\rho^{\backprime}$ of (1). Hence

$$d\rho^{\backprime} = \chi d\rho, \qquad d\dot{\rho}_b^{\backprime} = \chi d\dot{\rho}_b, \tag{5}$$

by the properties of extended vector linities.

When we say that the integrals (4) are the same as the integrals (3), we assert that each of the four elements of integration in (3) is equal to the corresponding element in (4). In other words, we assert the existence of four invariants,

$$\left.\begin{aligned} V_0\sigma d\rho = V_0\sigma^{\backprime} d\rho^{\backprime}, &\qquad V_0 d\dot{\rho}_2 V_2\nabla\sigma = V_0 d\rho_2^{\backprime} V_2\nabla^{\backprime}\sigma^{\backprime} \\ V_0 v d\dot{\rho}_b = V_0 v^{\backprime} d\dot{\rho}_b^{\backprime}, &\quad V_0 d\dot{\rho}_{b+1} V_{b+1}\nabla v = V_0 d\dot{\rho}_{b+1}^{\backprime} V_{b+1}\nabla^{\backprime} v^{\backprime} \end{aligned}\right\}. \tag{6}$$

From (5) and the fundamental property of the conjugate of χ these at once give

$$\left.\begin{aligned} \sigma^{\backprime} = \chi'^{-1}\sigma, &\qquad V_2\nabla^{\backprime}\sigma^{\backprime} = \chi'^{-1} V_2\nabla\sigma \\ v^{\backprime} = \chi'^{-1} v, &\quad V_{b+1}\nabla^{\backprime} v^{\backprime} = \chi'^{-1} V_{b+1}\nabla v \end{aligned}\right\}. \tag{7}$$

(5) and (7) suggest multenions of types r, s, such that

$$r' = \chi r, \qquad s' = \chi'^{-1}s, \tag{8}$$

r and s are multenion functions of ρ, and r' and s' are associated functions (say of ρ') which produce the same integrals when ρ' is taken as the independent variable in place of ρ [see the synopsis above]. When (8) gives the rule of the association of the functions,

r and s are said to be contravariant and covariant respectively.

It is well to place here the obvious consequence of (8)

$$\mathrm{V}_0 rs = \mathrm{V}_0 r's', \tag{9}$$

that is, $\mathrm{V}_0 pq$ is always an invariant if one of the two, p, q, is covariant and the other contravariant.

By exactly similar reasoning (a little complicated by the presence of k), we may treat of the equations (11) and (8) of § 8. For the sake of saving space, I merely state the results, k being given by (9) of § 8, and τ, u, $d\varsigma$ (including as particular cases $d\alpha$ and db) being as in (11) and (8) of § 8, the following are covariant:—

$$k\dot{\upsilon}, \qquad kdb, \qquad k\alpha, \qquad kd\varsigma_a$$

and the following are contravariant

$$kdb, \qquad \tau, \qquad u, \qquad k^{-1}\mathrm{V}_0\nabla(k\tau), \qquad k^{-1}\mathrm{V}_{a-1}\nabla(ku),$$

kdb is both covariant and contravariant. This merely means that it is invariant ($kdb = k'db'$). Remembering that $\chi h = h = \chi'^{-1}h$ when h is a scalar, (8) shows at once that, for a scalar covariance and contravariance, both mean precisely the same as invariance.

The persistent manner in which k thrusts itself into these results tempts one to wonder whether covariant and contravariant terminology is superior to intensity and flux terminology. After some experience of using both, I am inclined to believe that the accepted covariant and contravariant standard scheme is superior to the other.

It is not difficult to show that $k\dot{\upsilon}$ satisfies the test of (8) for covariance, that is

$$k'\dot{\upsilon} = \chi'^{-1}(k\dot{\upsilon}).$$

If in (2) and (3) of § 6 we read χ in place of ϕ, we get theorems closely connected with multiplication and composition in the absolute differential calculus. Let v_a, v_b be covariant multenions of homogeneities, a, b. Then

$$v_{a+b} = \mathrm{V}_{a+b}v_a v_b, \tag{9}$$

is covariant (multiplication). If u_a, u_b, are similar contravariants then

$$u_{a+b} = \mathrm{V}_{a+b}u_a u_b, \tag{10}$$

is contravariant (multiplication). If u is contravariant (homogeneity a) and v is covariant (homogeneity b), then

$$u_{a-b} = \mathrm{V}_{a-b}uv, \tag{11}$$

is contravariant (composition) when $a > b$. Also

$$v_{b-a} = \mathrm{V}_{b-a}uv, \tag{12}$$

is covariant (composition) when $a < b$. Lastly, when $a = b$, then each is invariant, which is a special case of (9). A special case arises when we put the larger of the two, a, b equal to n.

ξ, any multenion linity, is said to be covariant when it produces a covariant multenion from a contravariant, thus

$$s = \xi r, \qquad s' = \xi' r'. \tag{13}$$

Conversely since then

$$r = \xi^{-1}s, \qquad r' = \xi'^{-1}s', \tag{14}$$

ξ^{-1} is said to be contravariant. From (8) we at once get

$$\xi' = \chi'^{-1}\xi\chi'^{-1}, \qquad \xi'^{-1} = \chi\xi^{-1}\chi'. \tag{15}$$

From this, when ξ is covariant, its conjugate ξ' is also covariant, and, therefore, both the self-conjugate part and the skew part of ξ are covariant. Similarly for ξ^{-1} and contravariance. Also, when ξ is self-conjugate, ξ' is self-conjugate. If ξ is a covariant vector linity, the extended linity is also covariant. Also, in this case, the rotation-${}_n\mathrm{V}_2$ (see (10) of § 4) $\frac{1}{2}\mathrm{V}_2\zeta\xi\zeta$ is covariant. To prove this last, we have to show that

$$\chi'\mathrm{V}_2\zeta\xi'\zeta = \mathrm{V}_2\zeta\xi\zeta.$$

Now

$$\begin{aligned} \chi'\mathrm{V}_2\zeta\xi'\zeta &= \mathrm{V}_2\chi'\zeta\chi'\xi'\zeta && [(2)\ \S\,6] \\ &= \mathrm{V}_2\zeta\chi'\xi'\chi\zeta && [(9)\ \S\,5] \\ &= \mathrm{V}_2\zeta\xi\zeta && [(15)\text{ above}], \end{aligned}$$

as required.

Let $\sigma_1\ \sigma_2\ \ldots,\ \sigma_n$ be n independent covariant vectors. [They might be taken as ∇h_1, $\nabla h_2\ \ldots\ \nabla h_n$ where the Jacobian of the quantities h does not vanish, that is, where $\mathrm{V}_n(\nabla h)^{(n)} \neq 0$.] From the relations

$$\left.\begin{aligned} 1 &= \mathrm{V}_0\sigma_1\bar{\sigma}_1 = \mathrm{V}_0\sigma_2\bar{\sigma}_2 = \ldots \\ 0 &= \mathrm{V}_0\sigma_1\bar{\sigma}_2 = \mathrm{V}_0\sigma_1\bar{\sigma}_3 = \ldots \end{aligned}\right\}, \tag{16}$$

it at once follows that $\bar{\sigma}_1$, $\bar{\sigma}_2$, … $\bar{\sigma}_n$ are n independent contravariant vectors, and σ_c, $\bar{\sigma}_c$ might just as properly have been denoted by $\bar{\tau}_c$, τ_c. Any covariant or contravariant vector can be expressed at once in terms of these and invariants. Also, from the 2^n vectoriums formed from each set, any covariant or contravariant multenion may be similarly expressed. Thus the expressions for σ, τ, v, u are [see (2) of § 5]

$$\left.\begin{aligned} \sigma &= \sum_{c=1}^{n} \sigma_c \mathrm{V}_0 \sigma \bar{\sigma}_c, & \tau &= \sum_{c=1}^{n} \bar{\sigma}_c \mathrm{V}_0 \tau \sigma_c, \\ v &= \Sigma \mathrm{V}_b \sigma^{(b)} \,.\, \mathrm{V}_0 v \overline{\mathrm{V}_b \sigma^{(b)}}, & u &= \Sigma \overline{\mathrm{V}_a \sigma^{(a)}} \,.\, \mathrm{V}_0 u \mathrm{V}_a \sigma^{(a)} \end{aligned}\right\}. \qquad (17)$$

§ 10. *Differentiation.*—Our integration theorems have suggested certain systematisations, now to be given in the use of ∇. We have

$$\nabla w = \mathrm{V}_{c+1} \nabla w + \mathrm{V}_{c-1} \nabla w.$$

If w is covariant, one part on the right, viz., the first, is covariant and the other not. A similar statement applies to contravariance. We cannot, as seems desirable, resolve the symbolic vector ∇ into two parts, to which each of these two discordant parts on the right belong. But we can so resolve the symbolic *linity* $\nabla(\ )$. The first part of the linity, connected with $\mathrm{V}_{c+1}\nabla w$ produces a part of ∇w of homogeneity one unit *higher* than that of w, and may be thought of as the *ascending* derivative linity, and will be denoted by H. The other part, L, produces a *lower* homogeneity, and may be thought of as the *descending* derivative. Thus, formally, we define the linities H and L thus

$$\left.\begin{aligned} &\nabla(\ ) = \mathrm{H} + \mathrm{L} \\ \text{where}\quad &\mathrm{H} = \sum_{c=0}^{n-1} \mathrm{V}_{c+1} \,.\, \nabla \mathrm{V}_c(\ ), \qquad \mathrm{L} = \sum_{c=1}^{n} \mathrm{V}_{c-1} \,.\, \nabla \mathrm{V}_c(\ ) \end{aligned}\right\}. \qquad (1)$$

$$\left.\begin{aligned} \text{Otherwise}\quad & \nabla w = \mathrm{H}w + \mathrm{L}w \\ \text{where}\quad & \mathrm{H}w = \mathrm{V}_{c+1} \nabla w, \qquad \mathrm{L}w = \mathrm{V}_{c-1} \nabla w \end{aligned}\right\}. \qquad (2)$$

There are two, wholly separate, aspects of H and L as operators. Each is both a linear operator and a differential operator. In (2) the subject of these two operations is the one symbol, w, but in (4), below (where H′ is the conjugate of H), in

$$\phi_g \mathrm{H}_g' d\rho_{b+1},$$

we have a case where H′, as linear operator, affects the immediately succeeding symbol, $d\rho_{b+1}$, while as differential operator it affects ϕ, and this is indicated by suffixes in the usual manner. This example illustrates some of the facilities of operation provided by H.

In considering the fundamental properties of H and L, think of replacing

$$\begin{array}{lcccc} & \nabla, & \nabla w, & \mathrm{V}_{c+1} \nabla w, & \mathrm{V}_{c-1} \nabla w \\ \text{by} & \alpha, & \alpha w, & \mathrm{V}_{c+1} \alpha w, & \mathrm{V}_{c-1} \alpha w, \end{array}$$

where α is an ordinary vector instead of a symbolic vector. Let H′, L′

be the conjugates of H, L as usual. From the fundamental relation, $V_0 p\phi q = V_0 q\phi' p$, we at once have

$$H'w = V_{c-1} w\nabla, \qquad L'w = V_{c+1} w\nabla. \tag{3}$$

Thus the integration theorems (1) and (4) of § 8 read

$$\int^{(b)}\!\!\int \phi d\dot{\rho}_b = \int^{(b+1)}\!\!\int \phi_g H_g{}' d\dot{\rho}_{b+1}, \tag{4}$$

$$\int^{(n-a)}\!\!\int \psi ds_a = \int^{(n-a+1)}\!\!\int \psi_g L'_g ds_{a-1}, \tag{5}$$

and (7) and (8) of § 8 read

$$\int^{(b)}\!\!\int V_0 v d\dot{\rho}_b = \int^{(b+1)}\!\!\int V_0 d\dot{\rho}_{b+1} H v, \tag{6}$$

$$\int^{(n-a)}\!\!\int ukds_a = \int^{(n-a+1)}\!\!\int V_0 ds_{a-1} L(ku). \tag{7}$$

Further we have

$$H^2 = 0 = L^2, \tag{8}$$

and by taking conjugates we have

$$H'^2 = 0 = L'^2.$$

From (8) consider the various powers $\nabla^2(\ )$, $\nabla^3(\ )$ of $\nabla(\ )$. Clearly we get

$$\left.\begin{aligned} \nabla^2(\ ) &= HL+LH \\ \nabla^3(\ ) &= HLH+LHL, \text{ etc.} \end{aligned}\right\}, \tag{9}$$

and in (9) we may take conjugates [noting that, since the conjugate of $\nabla(\ )$ is $(\ )\nabla$, the conjugate of $\nabla^2(\ )$ is $(\ )\nabla^2$, etc.].

Notice that H′ *lowers* the homogeneity, while L′ *heightens* it, so that we have the following Table:—

	Heightens homogeneity.	Lowers homogeneity.
Prefixes ∇	H	L
Postfixes ∇	L′	H′

Either of the two, H and L, can be expressed in terms of the other thus

$$Lq = H(q\dot{v}) \,.\, \dot{v}^{-1}, \qquad Hq = L(q\dot{v}) \,.\, \dot{v}^{-1}. \tag{10}$$

Certain identities satisfied by H and L are given by the theorems of integration. As an example, $\int d\rho` = 0$ for a closed curve. This can be expressed in terms of ρ thus:

$$0 = \int \chi d\rho = \iint \chi_g V_1 d\dot{\rho}_2 \nabla_g.$$

Since the whole tract may be taken as the element $d\dot{\rho}_2$ we get

$$\chi_g V_1 d\dot{\rho}_2 \nabla_g = 0. \tag{11}$$

If we operate by $V_0\gamma(\ )$ we get

$$V_0 d\dot{\rho}_2 V_2 \nabla\chi'\gamma,$$

or since $d\dot{\rho}_2$ is arbitrary,

$$V_2\nabla\chi'\gamma = 0. \tag{12}$$

Both (11) and (12) are simple enough to be established directly from the form of χ, namely, $\chi\gamma = -V_0\gamma\nabla \cdot \rho^{\backprime}$, but the method of deriving them, just given, suggests how to prove other results not nearly so easy to establish directly. Instead of $\int d\rho^{\backprime} = 0$ we may take either of

$$\int^{(b)}\int d\dot{\rho}_b^{\backprime} = 0, \qquad \int^{(n-a)}\int ds_a^{\backprime} = 0,$$

and modify them in the same way. Thus from the first, $0 = \int^{(b+1)}\int \chi_g H_g' d\dot{\rho}_{b+1}$. From this and by taking conjugates we get

$$\chi_g H_g' q = 0, \qquad H_g\chi_g' q = 0, \tag{13}$$

where q is an arbitrary multenion. If we put $q = w$ (13) becomes

$$\chi_g V_{c-1} w \nabla_g = 0, \qquad V_{c+1}\nabla_g\chi_g' w = 0. \tag{14}$$

Similarly we have

$$0 = \int^{(n-a)}\int ds_a^{\backprime} = \int^{(n-a)}\int m\chi^{1-1} ds_a.$$

m is here $k/k^{\backprime}$, *i.e.*, it is the Jacobian of $\rho^{\backprime}$, or by (11) of § 7

$$m = n' V_0 V_n \nabla^{(n)} V_n \epsilon^{(n)}.$$

Thus

$$0 = \int^{(n-a+1)}\int (m\chi^{1-1})_g L_g' ds_{a-1}.$$

Hence, and by taking conjugates

$$(m\chi'^{-1})_g L_g' q = 0, \qquad L_g(m\chi^{-1})_g q = 0. \tag{15}$$

Putting $q = w$ we get

$$(m\chi'^{-1})_g V_{c+1} w \nabla_g = 0, \qquad V_{c-1}\nabla_g(m\chi^{-1})_g w = 0. \tag{16}$$

Let us now find expressions explicitly giving $\nabla^{\backprime}$, $H^{\backprime}$, $L^{\backprime}$ in terms involving ∇, H, L. By (2), § (9), ∇ is a symbolic covariant vector, so that

$$\nabla^{\backprime} = \chi'^{-1}\nabla, \tag{17}$$

where on the right the differentiations are not to affect χ'^{-1}.

For H we have that Hs is covariant when s is. Hence

$$H^{\backprime}s^{\backprime} = \chi'^{-1}Hs = \chi'^{-1}H(\chi's').$$

Now, by the second of (13) in this expression we may show the differentiations of H explicitly as affecting or as not affecting the immediately following χ^1, because the terms implied by those differentiations are identically zero. Hence we may write

$$H^{\backprime} = \chi'^{-1} . H . \chi', \tag{18}$$

bearing this permissibility of interpretation in mind. The relations between

H' and χ, on the one hand, and H' and χ^{-1}, on the other, are precisely the same, so that we may write

$$H = \chi' \,.\, H' \,.\, \chi'^{-1}, \tag{19}$$

with permissibility of interpretation as in (18).

Similarly from (18) and (19) by means of (5) of § 7

$$L' = (m^{-1}\chi) \,.\, L \,.\, (m\chi^{-1}), \tag{20}$$

$$L = (m\chi^{-1}) \,.\, L' \,.\, (m^{-1}\chi), \tag{21}$$

with the same permissible interpretation as to the incidence of the application of the differentiations of L and L' on the right; they may be supposed to affect the immediately following $(m\chi^{-1})$ and $(m^{-1}\chi)$ or not, at our pleasure.

A New Form of Wehnelt Interrupter.

By F. H. Newman, M.Sc., A.R.C.Sc.

(Communicated by Lord Rayleigh, F.R.S. Received February 21, 1921,—
Received in revised form, March 18, 1921.)

[Plates 3 and 4.]

1. *Introduction.*

The original form of Wehnelt Interrupter has many disadvantages. Unless special precautions for cooling the apparatus are taken, the solution soon boils, and the interrupter ceases to work. The large current density causes rapid disintegration of the platinum wire, and there is considerable expense in renewing it. The interrupter cannot be used with alternating currents owing to the melting of the wire when it is the cathode.

Sulphuric acid has a fairly large electrical conductivity, compared with other electrolytes, and the mean value of the current density at the wire electrode is large. There is, as a result, considerable heat developed in the volume of the acid. If a high resistance electrolyte is substituted, the current density at the platinum wire is much smaller, and the heating effect is reduced. The fumes arising from the acid, and also the spraying, when the interrupter is in action, are objectionable.

Although electrolytic interrupters are not widely used in the laboratory, they are in considerable demand for X-ray work, because of the heavy disruptive discharges which are obtained in the secondary of a coil operated by these interrupters. Providing special arrangements are made for the cooling, they require very little attention, and the interruptions are very steady and regular. The new form of Wehnelt Interrupter described below has many advantages over the older form.

2. *Description of New Form of Interrupter.*

The anode consists of a platinum wire which protrudes from a glass or porcelain tube. It is capable of adjustment by means of a threaded rod to which it is attached. A milled head is fixed to the rod, and so, by screwing up or down, the amount of the platinum surface exposed to the action of the electrolyte can be varied. In this way the platinum wire can be renewed without dismantling the apparatus; also the mean value of the current can be altered. The anode consists of an aluminium vessel which contains the electrolyte. The vessel thus serves a double purpose, and is more portable than a glass vessel. The electrolyte is a saturated solution of ammonium phosphate. If the solution is unsaturated, the interrupter works well, but requires a higher applied potential difference before the interruptions commence. This electrolyte has a very high specific resistance, and the current density at the wire is much lower than in the old form of Wehnelt Interrupter. As a result there is less local heating in the neighbourhood of the platinum, and less disintegration. Aluminium, with the ammonium phosphate, confer rectifying properties on the interrupter. The current will pass only from the wire to the aluminium in the solution, and, when alternating currents are employed, one-half of the wave is suppressed. Thus the wire never becomes the cathode, and there is no disintegration. No acid fumes are emitted, and even at 90° C. the interruptions are steady and regular. The older form of Wehnelt Interrupter ceases to act at this temperature. There is no fatigue of the platinum wire in this modified form.

3. *Experimental Results with the Modified Form.*

Table I shows the results obtained when the surface area of the wire exposed to the electrolyte was altered. The applied potential difference was that necessary to start the interruptions. The current increases as the surface area increases, but the current density remains constant. As the surface area increases, the applied potential difference necessary to start the interruptions also increases, and is proportional to the mean value of the current.

Table I.

Mean value of current.	Applied P.D.	Surface area of wire.	Current density.
ampères.	volts.	sq. mm.	ampères per sq. mm.
12·6	68	88	0·14
10·8	64	81	0·13
8·1	60	74	0·11
6·1	50	57	0·11
4·5	44	39	0·12
2·9	36	23	0·13
2·0	34	15	0·13
1·1	28	7	0·15

Table II indicates that the applied potential difference necessary to start the interruptions decreases as the temperature of the electrolyte increases. Providing the temperature is not too great—above 90° C.—it is advantageous to have the electrolyte hot. The mean value of the current decreases as the temperature rises, but this is due to the decrease in the applied potential difference.

Table II.

Temperature.	Applied P.D.	Mean value of current.
° C.	volts.	ampères.
18	52	6·6
32	44	6·4
45	34	6·2
65	26	6·1
69	23	6·0
78	18	5·7
93	16	5·6
96	15	5·5

4. *Wave Form of Primary Current.*

The wave form of the primary current was investigated by means of oscillograms. The oscillograph used was a Duddell permanent magnet type, having two vibrators specially stretched, so that they were able to follow the variations in the current up to 500 interruptions per second. The resulting curves were photographed by means of a falling plate camera, the speed of the plate at the instant of photograph being about 400 cm. per second. One of the vibrators—in series with a high non-inductive resistance—was placed in shunt with the main current; the other, with a non-inductive resistance in series, was connected across the terminals of the interrupter. In this way it was possible to observe on the same oscillogram, the form of the primary current, and also the variation of the potential difference across the interrupter. The resultant curves are shown in the Plate, the unequal brightness being due to the use of an alternating current arc as the source of light.

With direct currents the interruptions are extremely regular, both as regards frequency and the amount of current interrupted. The rate of decrease of the current at interruption is not uniform, but is very great at the commencement. As a result, the peak value of the potential difference across the terminals of the interrupter is very great—more than six times the mean value of the potential difference as measured on a direct reading voltmeter. The secondary discharge is thus made very disruptive. The ordinary gas bulb used for the production of X-rays, requires a secondary current of large peak value for good working. It is only the peak value of the current which is active, the rest of the current merely producing a heating effect in the bulb. The latter also works well with high frequencies, providing the frequency is not so great that the rate of make of the primary current becomes great. If this is the case, then there is a considerable reverse current in the bulb. Plate 3, I, shows that the time interval of break, compared with the time interval of make, is very small. There is little reverse secondary current. The peak value of the potential difference across the terminals of the interrupter is very great, even when the secondary current is flowing. This is the case also with large frequencies of interruption. The time interval of break is small compared with that at make, and, as a result, the secondary discharge is unidirectional at high frequencies. This is shown in No. II; not all of the current is interrupted, and the greater the frequency the less is the fraction of current interrupted. There is no trace of oscillations in the primary current at break, but, as the current never reaches zero, the conditions are against the detection of oscillations.

With alternating currents there is rectification, one half of the wave being practically suppressed. There is never complete interruption. The modified form works steadily without any self-induction in the circuit. Oscillations in the primary current of an electrolytic interrupter have not been observed previously, and it has been concluded that the gas of polarisation plays no part in the interruption of the current. Plate 3, III, shows decided ripples in the rectified and interrupted portions of the current. These oscillations are present only when there is no self-induction in the circuit. The introduction of self-inductance into the primary circuit changes the wave form of the primary current, increases the frequency of interruption, and decreases the fraction of the current interrupted. This is shown in IV. The secondary current affects the primary current, increasing the frequency of interruption, and decreasing the fraction of current interrupted, V. When the frequency is large—700 interruptions per second—the form of the primary current is irregular, and the interruptions are not equally distributed throughout the half-wave, being more frequent in the first half than in the second half (Plate 4, VI).

Although the loss of one half of the current by rectification means a loss of energy, the heating effect in the interrupter is reduced. When sulphuric acid is used, interruptions take place in both halves of the primary current. but the amount of current interrupted is very small, and varies in the two halves (VII).

As the secondary current in the coil was not great enough to produce a suitable deflection in the oscillograph, it was sent through an air transformer, and the transformed current passed through one of the vibrators. In this way, the wave form of the secondary current, produced under various conditions, was investigated. No. VIII shows the result obtained when a discharge tube was placed in the secondary circuit. The peak value of the current is great compared with the average value. There is practically no reverse current. If the old type of interrupter is used with alternating currents, the secondary discharge is also alternating. This is shown in IX. When the new form of interrupter is employed with alternating currents, the secondary discharge is practically unidirectional, as indicated in X, which represents the result obtained with the secondary current when it is passed through a gas X-ray bulb.

DESCRIPTION OF PLATES.

Plate 3.

The coil was 10-in. spark gap with variable primary. Induction about 1000 henries with a current of about 1 milliampère.

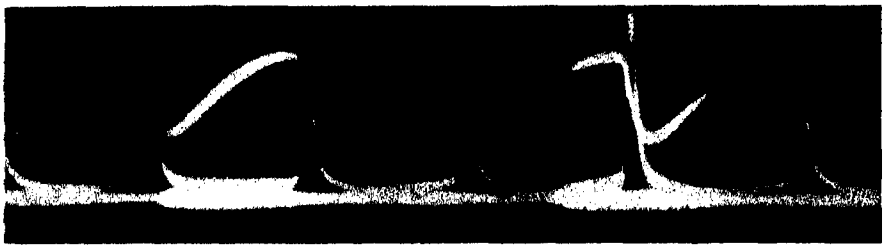
I. New form of interrupter. Direct current in primary. 60 volts, 7·4 ampères. No secondary current. Current and potential difference curves.

II. New form of interrupter. Direct current in primary. 60 volts, 7·0 ampères. Secondary current through spark gap 6-in. Current and potential difference curves.

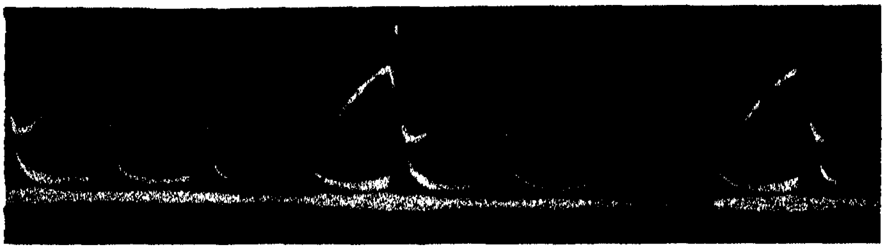
III. New form of interrupter. Alternating current through interrupter only. 100 volts, 4·0 ampères.

IV. New form of interrupter. Alternating current through primary of coil. 100 volts, 6·0 ampères. No secondary current.

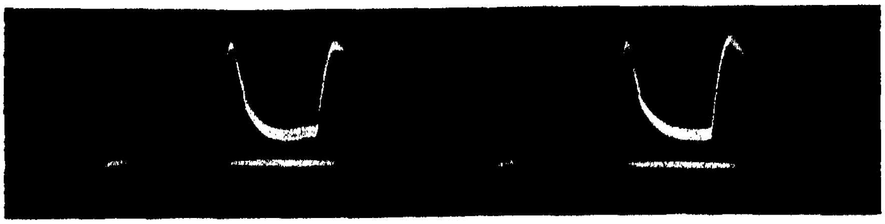
V. New form of interrupter. Alternating current, 100 volts, 7·0 ampères. Secondary current through 6-in. gap.



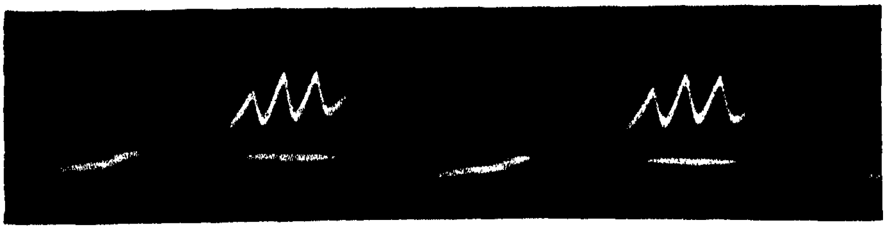
i



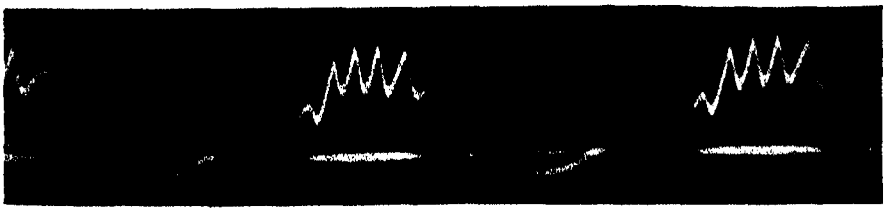
ii



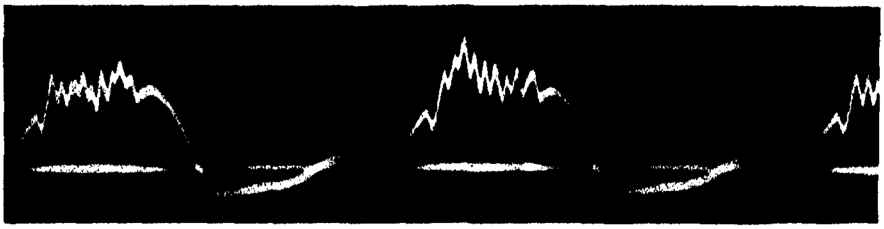
iii

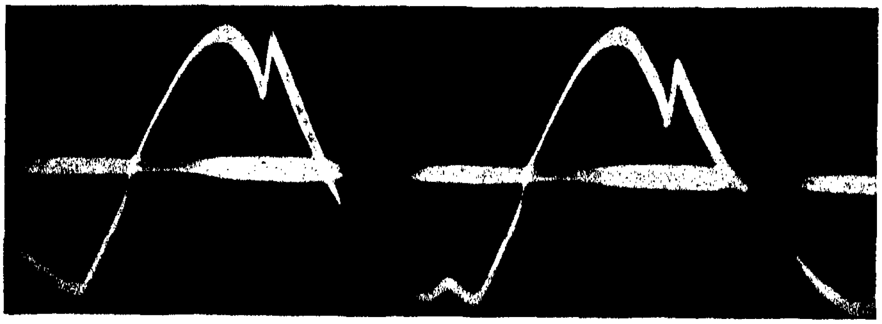


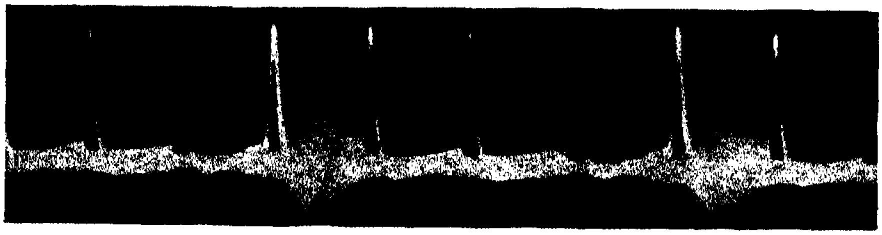
iv

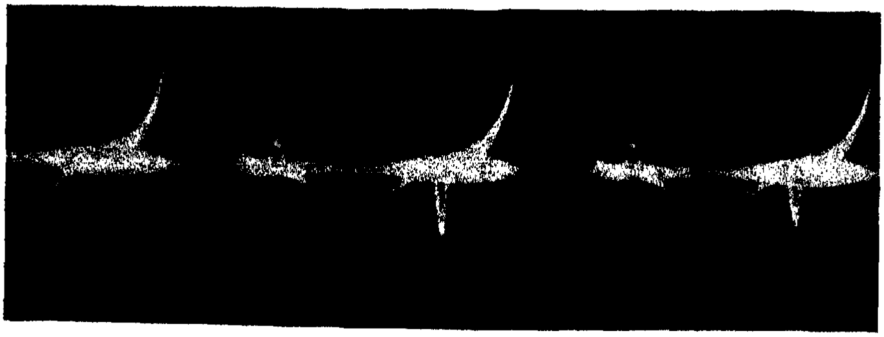


v


vi


vii


viii


ix

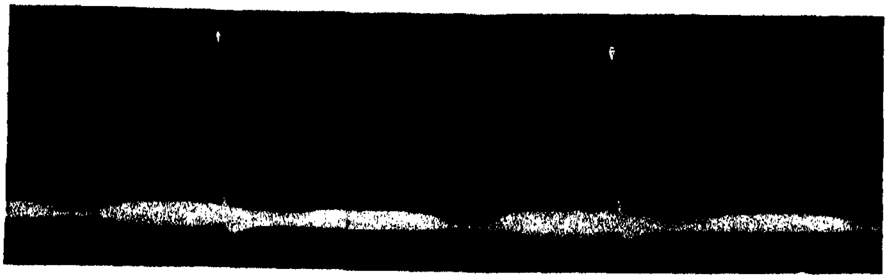

x

PLATE 4.

VI. New form of interrupter. Alternating current, 100 volts, 3·0 ampères. Secondary current through 4-in. gap.

VII. Old form of interrupter. Alternating current, 60 volts, 8·0 ampères. No secondary current.

VIII. New form of Interrupter. Curve of secondary current through discharge tube. Peak value of current 98 milliampères. Mean value 14 milliampères. Direct current in primary, 70 volts, 8·0 ampères.

IX. Old form of interrupter. Curve of secondary current through discharge tube. Alternating current in primary. 60 volts, 8·0 ampères.

X. New form of interrupter. Curve of secondary current through X-ray gas bulb. Alternating current in primary. 100 volts, 8·0 ampères. Peak value of secondary current 180 milliampères. Mean value 5 milliampères.

5. *Discussion of Results.*

The theory of the electrolytic interrupter is less known than that of the mechanical type. A relatively large resistance exists between the wire electrode and the electrolyte. This has been explained by Simon* on the supposition that the large current density causes vaporisation of the liquid in contact with the platinum wire, this layer of vapour forming the high resistance. Blondel† has indicated that the collection of oxygen gas, liberated by electrolysis, accounts for the resistance. He reproduced some oscillograph curves showing, that at 40 periods per second, the current is first established as in all inductive circuits. The electrolysis increases till a gaseous envelope is produced round the anode. Then there is a sudden interruption of the current, and, at the same time, an enormous increase in voltage. Then both the current and the pressure reassume their initial values, and the action begins agains. Blondel states that the energy $\frac{1}{2}LI^2$ accumulated in the self-induction, and liberated by the interruption of the current, causes the condenser formed by the polarised anode in the electrolyte to be charged at high potential. This condenser destroys itself by discharging in an arc which is formed between the platinum and the electrolyte, and drives away the dissociated oxygen. The oscillograph he used did not lend itself to an examination of the action at high frequencies but the oscillograms obtained in the present work, show that the phenomenon preserves the same character whatever the frequency. The value of the induced potential difference in the present experiments, as deduced from the oscillograms, amounts to 500–600 volts, so that the spark discharge can occur.

With alternating current, self-induction is not required in order to obtain the interruptions, and the fraction of the current interrupted is greater when self-induction is not present. Then, the disruptive force across the gaseous layer is small, and, as a result, the current remains steady, at interruption, for an appreciable interval before it is made again (III).

---

* Wiedemann, 'Annalen,' vol. 68 (1899).

† 'Comptes Rendus,' vol. 128 (1899).

The oscillograms show that the time interval of interruption is very small. Kallir and Eichberg* applied a stroboscopic method to determine the time interval of interruption, and found it to be of the order 0·0005 sec.; and making further stroboscopic observations, using the platinum-point first as anode and then as cathode, they found that the time interval with a point anode is about double that corresponding to a point cathode.

In the present experiments the gas, which was liberated at the anode, was collected, and found to consist of oxygen 80·2 per cent., and hydrogen 19·8 per cent. by volume. The disruptive spark which passes across the gaseous envelope, must cause some dissociation of the water vapour present, and this explains the presence of hydrogen. There is a copious volume of gas given off at the anode. If the insulating layer consisted of vapour, partial condensation would take place before these bubbles reached the surface of the liquid. This does not appear to occur.

It is probable that the heat effect and electrolysis together account for the production of interruptions.

6. *Summary.*

1. The new form of electrolytic interrupter consists of a platinum wire immersed in a saturated solution of ammonium phosphate. The whole is contained in an aluminium vessel, which sets as the cathode.

2. The current density at the anode is only one quarter of the value in the old form of Wehnelt Interrupter. Consequently there is less heating of the electrolyte, and less disintegration of the platinum wire.

3. The interrupter can be used with alternating currents, which it rectifies.

4. The secondary discharge, obtained with the new type of interrupter is very disruptive. It has a large peak value.

5. This modified type of interrupter works without self-induction in the circuit, when used with alternating currents.

6. The primary current wave form has been investigated with direct and alternating currents.

7. *Conclusion.*

Part of the cost of the apparatus used in the experiments described above has been defrayed by a Government grant, through the Royal Society, for which the author wishes to express his sincere thanks.

---

* 'Zeitschr. Electrotechn. Wien,' vol. 17 (1899.)

On the Viscosity and Molecular Dimensions of Gaseous Cyanogen.

By A. O. Rankine, D.Sc., Professor of Physics in the Imperial College of Science and Technology.

(Communicated by Prof. H. L. Callendar, F.R.S.—Received April 8, 1921.)

The molecule of cyanogen is of particular interest in the theory of molecular structure put forward by Lewis* and Langmuir.† Each cyanogen radicle (CN) is supposed to be practically identical in the arrangement of its outer electrons with the molecule of nitrogen (N_2), except that it lacks one electron. The junction of two CN radicles, to form C_2N_2, is conceived to be brought about by the sharing of one pair of outer electrons, so that the shape of the molecule may be regarded as nearly identical with that of two nitrogen molecules, with their centres at a certain distance apart, at present unknown. A test of the validity of this view is possible on the basis of the theory recently suggested by the author,‡ with reference to the collisions between non-spherical gaseous molecules, provided that the molecular dimensions of N_2 and C_2N_2 deduced from viscosity measurements are known. Reliable values for nitrogen are already available, but not for cyanogen. The viscosity measurements recorded in this paper supply this deficiency.

Experimental Determinations.

The cyanogen was prepared by heating well dried cyanide of mercury in an evacuated hard glass tube, the gas, as it was formed, being pumped off and collected over mercury. A volume of about 50 c.c. at atmospheric pressure was thus prepared, and further purification consisted of fractionating it at the temperature of −80° C. At this temperature cyanogen is solid, but it has an appreciable vapour pressure, so that, by pumping off the vapour and rejecting the early portions, possible impurities such as air and carbon dioxide were removed. The middle part of the fractionation—about 20 c.c.—was preserved over mercury for the viscosity measurements.

The method used for determining the viscosity has already been described in detail.§ Actually, the same apparatus was used as in the author's measurements of the viscosities of the rare inert gases. Only a small quantity of gas was required—about 6 c.c. at atmospheric pressure. The necessary observations are of a very simple character, consisting merely in comparing the times of descent of a mercury pellet between two marks in a vertical tube, when the apparatus is filled with different gases or with the same gas at different temperatures. For further details, the paper just referred to may be consulted. The experiments undertaken were (1) the determination of the ratio of the viscosities of cyanogen and air at atmospheric temperature; and (2) the comparison of the viscosities of cyanogen at steam temperature and at atmospheric temperature. From these results, the absolute data for air being accurately known, those for cyanogen can be deduced. Observations were therefore taken with the apparatus filled at atmospheric pressure alternately with air and cyanogen, both being dried with phosphorus pentoxide just previously to introduction. Each time of transpiration here recorded is the mean of a considerable number—at least ten—consistent amongst themselves to within 1 per cent., and this mean is considered to be reliable to 0·1 or 0·2 per cent. The mercury pellet was about 4 cm. long, and the same one was used throughout.

---

* G. N. Lewis, 'Journ. Amer. Chem. Soc.,' vol. 38, p. 762.

† I. Langmuir, 'Journ. Amer. Chem. Soc.,' vol. 41, p. 868.

‡ A. O. Rankine, 'Roy. Soc. Proc.,' A, vol. 98, p. 360.

§ A. O. Rankine, 'Roy. Soc. Proc.,' A, vol. 84, p. 181.

Temperature 15° C.

Times of transpiration in seconds.		
Cyanogen.	Air.	Ratio, cyanogen/air.
82·95	150·9	0·5497

This ratio of the times of transpiration is somewhat higher than the ratio of the viscosities, on account of the different amounts of slipping of the two gases over the walls of the capillary tube. The mode of applying this correction is indicated in my previous paper (*loc. cit.*), and need not be repeated here. The result obtained is

$$\text{At } 15^\circ \text{ C., } \frac{\eta_{Cy}}{\eta_{air}} = 0·548,$$

where η represents viscosity. At this temperature the viscosity of air is $1·799 \times 10^{-4}$ C.G.S. units, so that

$$\eta_{Cy} = 0·986 \times 10^{-4} \text{ C.G.S. units at } 15^\circ \text{ C.}$$

The times of transpiration for cyanogen at atmospheric and steam temperatures were as follows:

Cyanogen.

Times of transpiration in seconds.		Ratio, t_{100}/t_{15}.
15° C.	100° C.	
82·95	106·1	1·279

This ratio is approximately equal to the ratio of the viscosities at the two temperatures, but there are various small corrections necessary, including one for the expansion of the apparatus. These have been allowed for by making measurements of the time ratio for air over the same temperature interval, and comparing it with the correct viscosity ratio. Thus the time ratio for air was found to be 1·217. The true viscosity ratio is 1·220 for this temperature interval, so that the total correction is a positive one of 3 parts in 1200. Hence, for cyanogen

$$\frac{\eta_{100}}{\eta_{15}} = 1{\cdot}282,$$

which gives for the absolute value

$$\eta_{\mathrm{Cy}} = 1{\cdot}264 \times 10^{-4} \text{ C.G.S. units at } 100^\circ \text{ C.}$$

If we assume that Sutherland's law holds for the variation of viscosity with temperature, the values of Sutherland's constant (C) and the viscosity at 0° C. (η_0) can be calculated from the above data. The results thus obtained are

$$C = 280,$$

and $$\eta_0 = 0{\cdot}935 \times 10^{-4} \text{ C.G.S.}$$

The viscosity of cyanogen at 0° C. has been estimated previously from Graham's transpiration experiments in 1846.* The value given is $0{\cdot}95 \times 10^{-4}$ C.G.S. That now obtained is probably more reliable and correct to about 0·2 per cent. The same degree of accuracy cannot be claimed for the figure given for Sutherland's constant, owing to the comparatively small temperature interval to which the measurements were necessarily limited. In this case the probable error is of the order of ten units, or between 3 and 4 per cent.

Calculation of Molecular Dimensions.

A knowledge of both η_0 and C enables us to calculate by means of Chapman's† formula, modified somewhat in its interpretation according to recent suggestions of the author,‡ the mean area presented in collisions by a molecule of cyanogen in the gaseous state. This area proves to be

$$\bar{A} = 1{\cdot}31 \times 10^{-15} \text{ cm.}^2$$

and is probably accurate to about 2 per cent. It is improbable that the cyanogen molecule can be regarded as spherical, and the precise significance of $\bar{A}$ is not apparent unless some definite shape is assigned to the molecule. As mentioned earlier, there are reasons for supposing that C_2N_2 is nearly identical in shape with two linked nitrogen molecules. According to the views of both Lewis and Langmuir, the nitrogen molecule has a very compact arrangement of its outer electrons. Lewis supposes them to be arranged in close pairs at the five corners of two tetrahedrons with a common face; Langmuir assigns a special arrangement with eight outer electrons at the corners of a cube, as in neon and argon, and attributes the inertness of the nitrogen molecule to this disposition. Both these views, and particularly Langmuir's, make the nitrogen molecule comparable with the atoms of the inert monatomic gases. Accordingly, there would appear to be the same justification for regarding N_2 as behaving in collision like a hard elastic sphere as Chapman has shown to be the case for argon, kryptom, and xenon. The diameter of this sphere, obtained from viscosity measurements, is known, and is given by

---

* Kaye and Laby's Tables.

† S. Chapman, 'Phil. Trans.,' A, vol. 216, p. 279.

‡ A. O. Rankine, 'Roy. Soc. Proc.,' A, vol. 98, p. 181.

$$2\sigma = 3{\cdot}15 \times 10^{-8} \text{ cm.,}^{*}$$

whence

$$\pi\sigma^2 = 0{\cdot}78 \times 10^{-16} \text{ cm.,}$$

$\pi\sigma^2$ being the collision area for the nitrogen molecule, treated as spherical.

It will be seen that the mean collision area of the cyanogen molecule is greater than that of the nitrogen molecule in the proportion

$$\frac{\bar{A}}{\pi\sigma^2} = 1{\cdot}31/0{\cdot}78 = 1{\cdot}68.$$

Let us suppose, with Lewis and Langmuir, that a cyanogen molecule has the size and shape of two nitrogen molecules with their centres at a certain distance apart, and that it behaves in collision like a hard body formed by two overlapping hard spheres each of which has the kinetic properties of a nitrogen molecule. This body will present in collision a target of variable area, the mean value of which is determined by the magnitudes of the constituent spheres and their distance apart. The estimation of this mean area is precisely the same problem as that dealt with by the author in a former paper (*loc. cit.*) in which it was shown that the mean area in question ($\bar{A}$) is given in complete elliptic integrals by

$$\frac{\bar{A}}{\pi\sigma^2} = 1 + \frac{4}{3\pi}\left\{\left(\frac{1}{k}+k\right)\mathrm{E}\left(k, \frac{\pi}{2}\right) - \left(\frac{1}{k}-k\right)\mathrm{F}\left(k, \frac{\pi}{2}\right)\right\},$$

where σ is the radius of each of the constituent spheres, and k is the ratio d/σ, $2d$ being the distance apart of the centres of the spheres.

If we now identity $\bar{A}$ with the actual mean collision area found experimentally for the cyanogen molecule, and $\pi\sigma^2$ with that of the nitrogen molecule, we can calculate the distance apart of the centres, obtaining thus

---

* The value given by Chapman for 2σ is $3{\cdot}10 \times 10^{-8}$ cm. This has been corrected by adopting Millikan's latest value for the number of molecules per cubic centimetre, and by taking into account Kia Lok Yen's recent accurate re-determination of the viscosity of nitrogen ('Phil. Mag.,' vol. 38, p. 582).

$$2d = 2{\cdot}30 \times 10^{-8} \text{ cm.}$$

That is to say, from this point of view, we expect the two cyanogen radicles, each behaving as a nitrogen molecule as regards size and shape, to have their centres at this distance apart, when combined to form C_2N_2. Or, we may say that the diameter of the outer electron shell of the cyanogen radicle is $2{\cdot}30 \times 10^{-8}$ cm.

It is not yet possible to appeal to X-ray crystal measurements for confirmation of this estimate, for cyanide crystals have not hitherto been so examined. Such results will no doubt be obtainable by the same methods as W. L. Bragg* has used so successfully for measuring the dimensions of elementary atoms. It is, however, significant that the crystals of potassium cyanide and those of the potassium halides are usually stated to be isomorphous, and that, in addition, we find that KBr and KCN have nearly identical molecular volumes† —43·1 and 42·8 respectively. Thus if CN replaces Br there is no appreciable change of volume, and we may conclude, tentatively, that the cyanogen radicle and the bromine atom have the same size. Now Bragg's estimate of the diameter of the domain occupied in crystal structure by a bromine atom is $2{\cdot}38 \times 10^{-8}$ cm., and he assigns the value $2{\cdot}35 \times 10^{-8}$ cm. to the diameter of the outer electron shell of this atom. This is remarkably close to the above value $2{\cdot}30 \times 10^{-8}$ cm., which has been deduced for the distance apart of the centres of the two CN radicles in C_2N_2, *i.e.*, in circumstances which involve the sharing of a pair of outer electrons. Should X-ray examination of potassium cyanide prove it to be strictly isomorphous with potassium bromide, therefore, there will be good grounds for regarding the cyanogen radicle and the nitrogen molecule as bearing the relation to one another assumed in the Lewis-Langmuir theory.

An estimate would thus also be provided for the diameter of the outer electron shell of the nitrogen molecule, it being nearly the same as that of the kryptom atom, for the bromine and krypton atoms have the same size. The kinetic behaviour of the two gases confirms this view, since the diameters of the hard spheres to which the nitrogen molecule and the krypton atom are equivalent are $3{\cdot}15 \times 10^{-8}$ cm. and $3{\cdot}19 \times 10^{-8}$ cm., respectively. This is somewhat remarkable, when we consider how different is the number of electrons in the two cases. The nitrogen molecule has only 14, while the krypton atom has 36. It follows that the packing of the electrons must be much closer in the krypton atom than in the nitrogen molecule. A similar statement holds for the bromine and cyanogen molecules, in which the numbers of electrons are 70 and 26 respectively, yet the collision areas are nearly identical,—1·28 as compared with 1·31.

---

* W. L. Bragg, 'Phil. Mag.,' vol. 40, p. 169.

† Molecular volume here signifies the ratio of molecular weight to crystal density, or the volume per gramme-molecule.

I wish to express my appreciation of the valuable assistance given by Mr. C. J. Smith, of this Department, in the preparation and manipulation of the gas used in these experiments.

The Properties and Molecular Structure of Thin Films of Palmitic Acid on Water. Part I.

By N. K. Adam, M.A., Fellow of Trinity College, Cambridge.

(Communicated by W. B. Hardy, Sec.R.S. Received March 14, 1921.)

The experiments to be described in this paper have been undertaken in order to obtain confirmation of the very interesting views put forward by Langmuir (1) upon the arrangement of the molecules of various substances spread upon the surface of water. The study of these films has been carried on by a number of workers for many years, some of the principal publications being those of Miss Pockels (2), Rayleigh (3, 4), Hardy (5), Devaux (6), and recently Labrouste (7): but in most of these papers the authors do not enter into much detail regarding the molecular structure of the films.

So much information is now available however as to the dimensions of molecules and the forces about them, much of it being of an accurate quantitative nature, derived from structural organic chemistry, from the study of crystals, the kinetic theory of gases and the deviations from the simple gas laws, etc., that an attempt to deduce the arrangement of the constituent molecules from the properties of the films has become something more than a speculation and may be made with some certainty and definiteness.

And conversely, since the films of fatty acids and the other substances investigated appear to be only one molecule in thickness and to have all the molecules arranged in similar orientation—often simply perpendicular to the surface—measurement of the mechanical properties of films consisting of a known number of molecules of a pure substance, affords unusually direct information concerning the force fields round individual molecules. An accurate knowledge of these fields of force is of fundamental importance, not only to Physics but to practically all problems of Chemistry, including such as the nature of organic reactions, the theory of solution, and the questions of adsorption and colloidal stability.

Langmuir (*loc. cit.*) brings a great deal of evidence to show that the same force fields are the cause of chemical combination in its widest sense and of cohesion and surface-tension phenomena. This idea is confirmed in a remarkable manner by his experiments upon thin films of various insoluble fatty acids, esters, and alcohols, spread upon water. The molecules of such substances attach themselves to the water by those atomic groups which exhibit residual polar affinity in the reactions of organic chemistry. Thus the saturated fatty acids are attached by their carboxyl groups, the unsaturated acids having additional points of attachment at the ethylenic linkages in the hydrocarbon chain. This definite localization of the points of attraction in the molecule produces an orientation, the palmitic acid molecule turning perpendicular to the surface: while oleic acid, being attached by the carboxyl and also at a point in the middle of the chain, is bent down towards the water at the latter point.

Experimental Method.

The principle of the method of measuring the film's tendency to spread on water is the same as that first used by Miss Pockels, and later by Rayleigh and Hardy. A small measured quantity of the substance dissolved in a volatile solvent is placed on the cleaned surface of water: this spreads at once and evaporates in a few seconds, leaving a film behind. The area available for the film is now contracted gradually by moving barriers along the surface of the water, and the forces on the film are measured.

The compressive force on the film, *i.e.*, the difference between the surface tension of the film and that of pure water, has been measured directly by means of an apparatus similar to Langmuir's (*loc. cit.*, fig. 6), a thin strip of paraffined copper being used as the floating barrier instead of paper. The force on this float was measured by a simple balance. The clearance between the ends of the float and the sides of the trough was 1–2 mm., and escape of the film past the float was prevented by air blasts from two fine jets carefully adjusted in position, as in Langmuir's experiments. These air blasts did not affect the film much except for a very small area near the float, nor was the balance disturbed except for a slight alteration of zero.

I have used an accurately made brass trough measuring 60 by 14·00 by 1·8 cm., the sides being nearly 1 cm. thick and machined and scraped flat on the top. The barriers used for confining the film were strips of window glass about

30 by 1·5 by 0·7 cm.: these rested on the sides of the trough and projected some distance so that they could be handled with little risk of grease spreading from the fingers to the water surface. The trough was filled to the brim and levelled on a separate table.

Precautions necessary in Cleaning and Using the Trough.

It was cleaned with fine emery cloth under the tap until it wetted all over: the barriers were kept in a basin of hot chromic acid, being removed with tongs and washed with hot water so as to dry at once. The trough was dried in a warm place and the top of the sides coated with paraffin wax dissolved in benzene: the under side of the barriers was also coated with wax.

A simple trough of tin plate with paper barriers similar to that used by Langmuir was tried, but the barriers in my hands proved leaky and troublesome to keep in position: the above arrangement appeared capable of giving more accurate results. Water will spread over a clean metal surface and it is therefore difficult or impossible to obtain barriers quite free from leaks if glass and metal surfaces alone are used; the object of the wax was to prevent spreading, and it was found to be quite effective (with careful working) in preventing leaks. Films have been maintained unchanged in area for several hours, under compression.

Accidental leakage occurred occasionally however and was generally due to one of the following causes: the air blasts were not properly adjusted or the edges of the float had become wet so that the effective channels past the ends of the float were more than 2 mm. wide; leaks at the glass barrier were more frequent however and were usually due to the paraffin wax having been partially rubbed off.

Freshly paraffined surfaces rarely gave any trouble.

Residual Contamination.

Perfectly clean water surfaces were not obtained: on the cleanest the float would move away from the barrier at 0·5 cm. distance, after the barrier had been moved from about 35 cm., with a weight in the balance pan of 0·05 grm. Part of this apparent contamination was due to the capillary curvature of the liquid surface in contact with the barriers. Surfaces with more than 1·2 cm. of contamination have not been used.

A clean atmosphere is essential for these experiments: work next door to a laboratory where much organic preparation was being carried on was found impossible, owing to the rapidity with which the surfaces became contaminated.

A correction was made for the contamination as follows: A blank experiment, made immediately prior to putting on the film, gave the total amount of contamination at 0·05 grm., and the compression curve obtained with the film which had accumulated when a clean surface was exposed to the air for some time gave the ratio of the area filled by the contamination at any force to the area at 0·05 grm.

This area was then subtracted from the observed area of a film of palmitic acid. Table I gives some corrections for a commonly occurring amount of contamination.

Table I.—Corrections to Area of Film for Residual Contamination.

Grm. in pan.	*Correction* (cm. length of trough).	Grm. in pan.	*Correction* (cm. length of trough).
0·05	0·75	0·5	0·25
0·1	0·63	0·75	0·15
0·2	0·43	1·0	0·08
0·35	0·33		

The amount of contamination is small, but, owing to its much greater compressibility, the correction to the slope of curves of palmitic acid is appreciable. The films used were usually between 20 cm. and 30 cm. in length.

About five drops of benzene were used in putting on a film, and the residue left by this quantity was negligibly small, being certainly not greater than 0·16 cm., and probably much less even than this. The benzene was free from thiophene, was distilled and collected over ½° C., and was finally purified from grease by redistillation in an apparatus free from cork or rubber.

The water was distilled from tap-water in a large copper vessel and condensed in glass, all joints in the still being either ground or cemented with sodium silicate and plaster of Paris.

The palmitic acid was Kahlbaum's K quality, recrystallised from aqueous alcohol. It melted at 62·5° C. The curves show its freedom from unsaturated substances.

The quantity of solution put on the surfaces was measured by counting drops from a capillary pipette. The method is accurate within 3 to 5 per cent., and seems to be much more exact than any other simple method of measuring the small volumes.

The acidity of the water has been controlled by dissolving "buffer" salts in it, the concentration never exceeding M/65, which was generally enough to maintain the P_H* constant within 0·2 unit or nearer throughout a day's exposure.

---

* P_H of x denotes a hydrogen ion concentration of 10^{-x}.

The actual P_H was determined by colorimetric comparison with the standards of Clark and Lubs (8), which were checked against a hydrogen electrode.

Comparison experiments on distilled water indicated that these salts in the dilutions employed had little, if any, effect on the films, beyond regulating the hydrogen ion concentration.

$6{\cdot}06 \times 10^{-23}$ has been taken as the Avogadro constant.

Range of Temperature.

The profound change in the films with rising temperature described by Labrouste (7) was ascertained not to occur on any of the solutions investigated until the temperature had risen well above 20° C.

No measurable change occurred between 10° and 20°, and all the experiments described were done within this range, except where specially mentioned.

Discussion of the Curves.

Fig. 1 gives the uncorrected, I, and corrected, II, curves relating the force of compression to the area per molecule on solutions having a P_H between about 8 and 5·5. At areas greater than about 22×10^{-16} sq. cm. per molecule there was no measurable resistance to compression. At this point there was a sudden opposition to further contraction, and the compressive force increased practically linearly with decreasing area. The slope of this line gives the force required for 1 per cent. decrease in area as 5·9 dynes per centimetre: if the density of the material of the film is unity, its thickness is $2{\cdot}11 \times 10^{-7}$ cm., and the force required for 1 per cent. contraction, calculated on 1 sq. cm. cross-section, is 28 megabars. This slope is constant in successive experiments within about 15 per cent., which is probably the error of experiment. It is considerably steeper than the slope of the curves given by Langmuir (*loc. cit.*, fig. 8) for palmitic acid.

The value of the compressibility of the paraffin $C_{16}H_{34}$ obtained from Landolt Börnstein's tables indicates that a force of 125 megabars is required to compress this substance by 1 per cent. of its volume. This force is about 4·5 times as great as for the thin film, the latter being free to expand in thickness during compression.

The co-ordinates of the point H, at which collapse of the film was first detected, were not found susceptible of accurate determination, values from 20 to 45 dynes per centimetre being obtained. There are 95 and 214 megabars respectively. It appeared that the breakdown of a film might be delayed considerably beyond the point at which there would be equilibrium between a film and the substance in bulk on the water. A discussion of the probable reason for this metastable condition will be given later. The values for the force at H are much lower than those given by Langmuir.

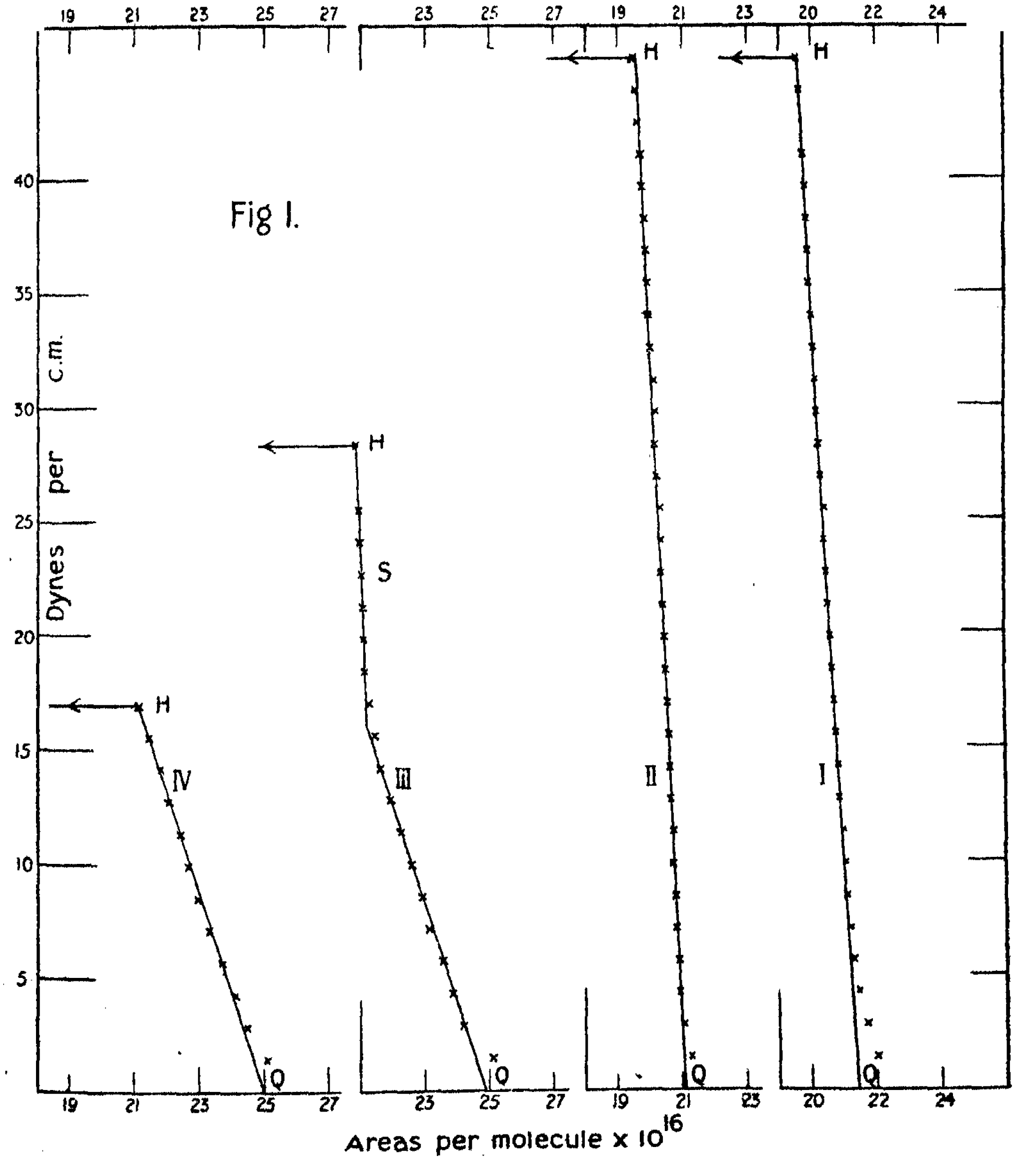


Fig I.

From the considerations advanced by Hardy ((5), 1912, p. 319), the probable value for the compression at which a film is in equilibrium with the acid in bulk resting on the surface is about 25 dynes per centimetre near the neutral point.

Effect of Acid. (Fig. 1, III and IV.)

When the hydrion concentration exceeded a point about $P_H = 5{\cdot}5$, the corrected curves became two straight lines, the upper portion being coincident within the error of experiment with the curve on neutral solutions, and the lower being of less steep slope and cutting the axis of no force at about $25{\cdot}2 \times 10^{-16}$ sq. cm. The intersection of the two lines was at about 15 dynes per centimetre at P_H of 5. The change of zero area was completed within a narrow range of P_H; but further increase of acid raised the point of intersection slightly, to about 19 dynes per centimetre, at a P_H of 3.

Physical State of the Films.

Motes on the surface gave (as has been pointed out by Langmuir and Labrouste) a means of seeing whether the films were solid or liquid. In the solid films motes were not blown about by the air blast: in liquid films, motes moved about without difficulty.

Solidification occurred usually as soon as the film was formed on neutral solutions (at Q) and on acid solutions at or soon after the steep part of the curve was reached. Occasionally, however, there was some delay in solidification: this did not appear to affect the compressibility. In the steep part of the curves, apparently, the molecules are arranged so that they may form a solid film, though this may sometimes be delayed; in the less steep part the molecules are not capable of forming a rigid film. The probable difference in the arrangement of the molecules will be discussed later.

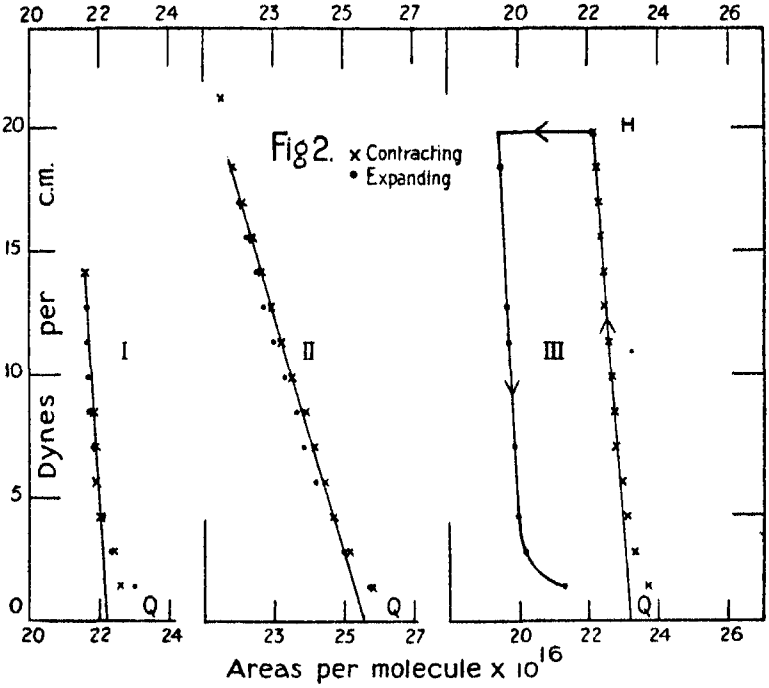
Absence of Hysteresis.

If the films below H are of 1 molecule in thickness, then there should be no hysteresis until collapse has set in. This was very nearly the case for the curves on neutral solutions and for the less steep portion on acid solutions. (Fig. 2, I and II.)

Films have also frequently been left under a compressive force for a long time, without change of area or any noticeable modification of properties. In one experiment the force was increased rapidly up to 15·5 dynes per centimetre, the film then left to itself for five hours, and the experiment then completed. The curve showed no discontinuity, although a change of 0·1 unit of area could have been detected.

This stability of the films showed the complete insolubility of palmitic acid in water; in the experiment just described not ½ per cent. of the molecules left the film, although each one was in contact with an enormous excess of water.

The only sign of hysteresis noticed before the point H was reached was that when the compression was increased above about 20 dynes per centimetre the points on the curve of expansion sometimes fell at slightly smaller areas. The difference may have been due to the residual contamination in the film.



Influence of Dust Settling.

Providing the film was under compression about 14 dynes per centimetre, contamination settling from the air had no effect on its area. At pressures much lower than this there was a slow expansion.

This is of some importance, since it shows under what conditions the use of lycopodium or talc sprinkled on the surface may cause error, unless special precautions are taken to purify the powder from oily matter.

Form of Curves at very Low Compressions.

The linear relation holds accurately down to a force of 3 or 4 dynes per centimetre, but it is uncertain whether the film occupies a greater area of still lower compressions. Langmuir (footnote, p. 1889) indicates that the curves *may* become asymptotic to the axis; and most of my experiments showed a slight divergence from the straight lines at 1·4 dynes per centimetre. The amount of divergence could not be measured with certainty, however, and it seems doubtful if the method is yet refined enough to decide whether the effect is a property of the palmitic acid or is due to the correction for residual contamination being imperfectly known, or to traces of impurity in the acid.

Table II gives the area occupied by the film at no compression under the head "Corrected area," assuming the linear relation to hold throughout, at different concentrations of hydrion. Column 2 gives the observed areas per molecule without any correction for residual contamination, and column 4 the salts used to regulate the hydrion concentration. Each figure is the mean of four or more observations. The areas are believed to be accurate within 5 per cent.

Table II.—Areas of the Uncompressed Films.

P_H.	Observed area.	Corrected area.	Salts in solution usually *about* M/70.
1	26·1	25·1	N/10 HCl.
2·9 to 3·1	26·3	25·3	KH phthalate + HCl.
4·2 4·4	27·2*	25·4	KH phthalate + NaOH.
5 5·2	25·8	25·1	KH phthalate + NaOH.
6	23·1	22·1	KH_2PO_4 + NaOH.
7	22·2	21·2	KH_2PO_4 + NaOH.
8·5	21·6†	21	H_3BO_3, KCl, NaOH.

* The curves on phthalate solutions at P_H 4 diverged considerably from a straight line near the base, for a reason unknown.

† At 8·5 there were signs of solubility of the films, which would diminish the observed area.

Structure of the Films.

The theory of Langmuir and Devaux that the films are one molecule in thickness between Q and H, each molecule being oriented with the carboxyl group towards the water and the hydrocarbon chain normal to the surface, is completely confirmed by these experiments. Ample evidence of this structure has already been given by Langmuir (*loc. cit.*) and Harkins and his collaborators (9). It may be emphasised, however, that the observed area of about 22×10^{-16} sq. cm., which is nearly equal to the estimated cross section of the hydrocarbon chain, is very much too small to allow any considerable number of the molecules to be arranged otherwise than vertically in the film: while considerations of solubility indicate that the carboxyl group will be directed towards the water rather than the other end of the molecule.

The slight diminution in area between Q and H, with a compressibility 4·5 times that of paraffin in bulk, is consistent with the theory that the molecules are being forced somewhat closer together, as in a liquid under compression. Stereochemical considerations indicate that the carbon atoms are not vertically above one another, but are probably arranged helically, the angle between two successive valency linkages being normally 109°, but being capable of increase or decrease under stress. In the film, the chains may be straightened out, while in the compression of paraffin in bulk, the pressure is applied also to the ends of the chains and prevents elongation, so that the rather greater compressibility in the case of the film is not inconsistent with the view that the chains are being crowded laterally without other alteration in position.

For other substances also, the evidence in favour of the layer being one molecule thick rather than two, appears to be conclusive, although Marcelin (10), following a suggestion of Rayleigh (4), contends that a double layer is formed. One of the principal arguments for the existence of a second layer has been that the area at H is for many substances, such as oleic acid, about half the area at Q. This great apparent compressibility of a unimolecular film does not necessarily mean however that the molecules are being reduced in volume by one half, which would indicate a true compressibility so high as to be out of the question, but it can be explained by the molecules being attached to the water by one point at H and by two points at Q, an explanation which is beautifully supported by Langmuir's discovery that the ratio of areas at Q and H bears an extremely simple relation to the number of points of residual affinity in the molecule. Thus, for oleic acid, with two points of affinity, the ratio is 2 to 1, for saturated acids it is very much less, and for ricinoleic acid, which has three points or more, the ratio is much greater: and, further, the areas were found to be approximately the same for all these substances at H. Clearly, the explanation is that the film is one molecule thick throughout, but that the molecules occupy an area roughly proportional to the number of points in contact with the water, being finally attached probably by the carboxyl group only.

Other reasons also negative the possibility of a second layer. There is evidence, both from the kinetic theory of gases and from recent very beautiful experiments by Langmuir (11), (12), on condensation, evaporation, and adsorption, that the attractive forces about the molecule extend to a range of a very few Ångström units only. Since the upper limit of the first layer is probably about 20×10^{-8} cm. from the water, any molecules in the second layer will be well out of range of the attraction, and will therefore not affect the surface tension of the water. And even if the molecular attraction is supposed to extend farther than this, it would be necessary to postulate (i) that the effect of a molecule in the first layer is *nil* until the molecules in this layer are repelling one another; (ii) while, on the other hand, molecules in the second layer lower the surface tension in exact proportion to the number present; and (iii) that the effect of molecules beyond the second layer becomes suddenly negligible. Some such postulates seem to be unavoidable, and it is difficult to see how they could *all* be simultaneously consistent with any imaginable law of force round the molecule: nor does it seem probable that modifications, which could be introduced in the conception of the film by considering the thermal motions, would much simplify the law of force needed to account for a bimolecular layer.

Probably, however, little error is involved in neglecting the thermal movements. There is a fairly complete analogy between the films at the temperature of these experiments and a solid. The films frequently possessed rigidity, and the properties were found to be unchanged, under certain conditions, over a range of 30° C. But at a well defined temperature, as Labrouste has shown, they "melt," undergoing a considerable increase in area and becoming mobile. In solids, the space lattice can be determined without taking account of the vibrations, and it seems probable that the thermal movements (in a vertical direction at least) are so small that the relative positions of the long molecules is not thereby disturbed.

Theory of the Transformation on Acid Solutions.

The following is suggested tentatively. From the data of molecular volume studies, the approximate cross-section of the palmitic acid molecule may be estimated. The volume of CH_2 is 17·8 c.c. at the m.p. of the paraffins; dividing by the Avogadro number and the distance, $1·52 \times 10^{-8}$ cm., between the carbon atoms in the diamond, which may be expected to be nearly the same as the distance between successive carbon atoms in a chain, since the linkages are of the same type, the cross section of the chain is found to be $19·3 \times 10^{-16}$ sq. cm. This is in fair agreement with the observed area occupied by the acid on neutral solutions. It is therefore probable that the steep parts of the curves indicate a repulsion between hydrocarbon chains not $-COOH$ groups.

A COOH group has a molecular volume about 1·7 times that of a CH_2 group, according to the data collected by Le Bas ((13), pp. 18 and 140). The length of the $-COOH$ group in the direction of the chain is unknown, and therefore its cross section is uncertain: it is probably, however, somewhat larger than that of the hydrocarbon chains.

The greater area on acid solutions will then be due to repulsion occurring between COOH groups: the smaller area on neutral solutions to these groups being so arranged that the hydrocarbon chains can come into close contact.

Probably this arrangement is simply that adjacent molecules are immersed in the water to different depths,* so that the carboxyls do not lie in the same horizontal plane. On acid solutions the molecules rise in the water till all are at approximately the same depth, and therefore the carboxyls touch, the chains being forced farther apart so that they cannot interlock to give a solid film.† A strong lateral compression (as in the upper part of the curves on acid solutions) will naturally force the carboxyls down so that there is again close packing.

The chains pack closely on neutral solutions without external lateral compression because the attraction of water for $-COOH$ groups increases with decrease in the hydrion concentration. There is independent evidence for this variation in attraction. The interfacial tension between water and fatty substances is less, the more alkaline the water, as was shown by Clowes (14) with strongly alkaline solutions, and recently by Peters and Hartridge (15) working near the neutral point. Oleic acid behaves similarly, and Dr. Peters informs me that benzene solutions of fatty acids also show the same phenomenon, which appears to be general for surfaces consisting of a layer of carboxyl groups oriented towards the water. It follows, from a well-known theorem of Dupré (see Hardy (5), 1912, p. 611), that the work gained when a layer of COOH groups approaches the water is greater for alkaline than for acid solutions. The chemical reaction between alkali and COOH would lead one to expect an effect in this direction.

The resistance of the hydrocarbon chains to immersion is clearly shown in the decreasing solubilities and increasing adsorption of fatty acids as the chains are lengthened: in any given acid then it is probable that increased attraction on the carboxyl groups will cause immersion to greater depths.

Solubility in Alkaline Solutions.

As the alkalinity increases, many of the chains become further immersed, though there is not much further alteration of the film's properties, since the chains are already closely packed at P_H 5·5; finally, the point is reached at which some molecules are completely immersed and solution commences. At this point the arrangement of minimum potential energy, in which the attractions are as completely satisfied as possible, will be that of aggregates having the carboxyl group outwards and the hydrocarbon chains packed in the centre. These will have a high molecular weight, and will probably carry many negative charges owing to the dissociation of the carboxyl groups.

---

* This idea was suggested to me in conversation by Dr. Peters.

† More recent experiments on other substances indicate that close contact of long hydrocarbon chains is probably not a necessary or sufficient condition for solidity of the films.

The genesis of the "ionic micelles" in aqueous soap solutions, which have been recently studied by McBain and his collaborators (16), thus becomes clear.

The film dissolved so fast that measurements of compression were impossible, at a P_H of 10; at 8·5 the curve showed signs of solubility, and at 8 (at room temperature), solution was too slow to measure. I hope to obtain detailed measurements of the rates of solution under various conditions later.

Influence of Acidity on the Melting Point of the Films.

The effect of rising temperature is being studied, and the fact noted by Labrouste (*loc. cit.*) that at a fairly well-defined temperature the films increase greatly in area under low compressions, and (if previously solid) become liquid, has been confirmed. It is striking, however, that a much higher temperature is required to melt the films on alkaline solutions than on acid, a fact which confirms the theory that the molecules are attracted more firmly to the water in the former case. Table III gives the approximate temperatures at which the first change occurs in the films.

Table III.

P_H.	Approximate m.p.	P_H.	Approximate m.p.
(1)	(25° C.)	(6)	(38°)
3	25°	(6·5)	(40°)
(5)	(33°)	8	Above 50°
(5·5)	(33°)		

Figures in brackets are provisional only.

Nature of the Collapse at H.

Fig. 2, III, shows the hysteresis effect observed after the collapse had proceeded a short distance. The curve of expansion was approximately parallel to the curve of compression of the original undamaged film, and this considered together with the facts that solid particles become visible after it is passed, and that solid palmitic acid does not spread upon water, indicates that the mechanism of the collapse was the destruction of a part of the surface film, molecules being ejected to form aggregates of relatively great thickness, the remainder of the film being unaltered in properties.

The solid films most frequently collapsed in the manner described by

Labrouste (*loc. cit.*), with the appearance of solid material in the form of very fine lines on the surface, roughly, arcs of circles extending from one end of the barrier to the other. Occasionally, however, the solid appeared in flakes, not lines, and when this occurred the flakes were usually, though not invariably, close to the air blasts.

Collapse, once started, sometimes proceeded to very small areas without further weights being required, the glass barrier being moved up slowly so as to keep the balance float nearly at zero. More often there were checks, frequently several in one experiment at irregular intervals, to overcome which the forces had to be increased by an amount indefinite, but occasionally nearly as high as the force given for H by Langmuir. When complete checks occurred, it was often possible to increase the force by a considerable amount before collapse set in again: and in these cases the curve would rise at the check at about the same slope as the curve of the original film, as would be expected if the resistance were due to an intact portion of the film with its original properties.

As might be expected from the diminished attraction on the $-COOH$ groups, the *minimum* force at H observed on acid solutions was less than that for neutral solutions, being sometimes as low as 16 dynes per centimetre (fig. 1, IV).

The nature of the collapse is easily accounted for. A single molecule leaving the film must overcome the attraction of the carboxyl group for the water, and also that of the hydrocarbon chain for its neighbours. The surface tension of paraffins and their solubility in one another indicates that this attraction is important.

If, however, a group of adjacent molecules are ejected together, the only work done is that of the carboxyl groups against the attraction of the water. Hence the film will collapse at comparatively few points only, and the solid flakes or lines on the surface are to be expected.

The metastable state and the checks to collapse are explained probably by the necessity of some nucleus for starting expulsion of the molecules at any point. Possibly these are dust particles, which have fallen on the surface before, or during, the application of the film, or at very low compressions, and have become incorporated with the film. The checks indicate that a nucleus in one part of a solid film may be useless for more than a certain number of molecules, being very likely expelled along with them.

Explanation of Langmuir's Curves on Acid Solutions.

The very sharp rise in resistance to compression shown at areas from ·2 to 8×10^{-16} sq. cm. per molecule, is, I am sure, only an instance of a similar check to collapse, and is due to the same cause (*loc. cit.*, figs. 10, 11, 12, 17). The areas at which these occur are much smaller than the cross section of any part of the molecules, and the effect is therefore very unlikely to be a property of an entire unimolecular film in any condition; and moreover, in numerous experiments, I have not observed any regularity in the phenomenon. It sometimes occurs on neutral solutions, as well as on acid; it does not invariably occur even on the latter; it may occur at any area after solidification of the film; and frequently there are several partial checks, at different areas, in one experiment.

The checks may be due to exhaustion of the nuclei available in the film, so that the remaining parts of the film, having no nuclei or less effective ones, are more resistant to collapse.

I should like to express my thanks to Mr. W. B. Hardy, F.R.S., for valuable advice during the progress of this work, and to Miss Jordan Lloyd for very kindly doing the electrometric determinations. The experiments were done in the Chemical and Biochemical Laboratories, Cambridge, and are being continued.

Summary.

1. A modification of the technique for examining the properties of thin films of substances on water is described, special attention being given to the avoidance of leaks past the confining barriers and to correcting for the residual contamination on the water which cannot be wholly avoided.

2. Experiments on palmitic acid confirm Langmuir's results in general. The curves of compressibility of the films are accurately straight lines, except for a possible doubtful deviation at very low compressions.

3. The film changes abruptly in properties at a hydrogen ion concentration about 10^{-5} N. The change is confined to films under compression less than about 16 dynes per cm. The area occupied by each molecule becomes about 20 per cent. greater than on neutral solutions.

4. Langmuir's view that the films consist of a single layer of molecules, oriented with their carboxyl groups towards the water, is confirmed and extended. Probably the hydrocarbon chains are tightly packed on neutral solutions, adjacent molecules being drawn into the water to different depths. On neutral solutions the molecules rise in the water so that the $-COOH$ groups, which are probably of greater cross section than the chains, are in contact with each other. This arrangement gives a liquid film, whereas the packed hydrocarbon chains normally give a solid one. There is independent evidence of an increased attraction of water for $-COOH$ groups with increasing alkalinity.

5. The formation of the "ionic micelle" of soap, as the film dissolves in more alkaline liquids, is a simple consequence of the increased attraction on the carboxyl groups and the attraction of the chains for each other.

6. Hysteresis in the unimolecular films of palmitic acid is absent or extremely small in amount, at room temperature.

7. Palmitic acid is completely insoluble in water slightly on the acid side of neutrality, the films preserving their area and properties unchanged for hours.

8. Compression beyond a certain point causes collapse of the film: the ejected molecules forming aggregates visible to the naked eye. The remainder of the film appears to be unchanged in properties. Different films, and even different parts of the same film, vary very much in their resistance to collapse. The collapse seems to be aided by nuclei, the nature of which can only be conjectured at present.

REFERENCES.

(1) Langmuir, 'Journ. Amer. Chem. Soc.,' vol. 39, p. 1868 (1917).
(2) Pockels, 'Nature,' vol. 43, p. 437 (1891).
(3) Rayleigh, 'Roy. Soc. Proc.,' vol. 47, pp. 281, 364 (1890); and vol. 48, p. 127.
(4) Rayleigh, 'Phil. Mag.,' vol. 48, p. 331 (1899).
(5) Hardy, 'Roy. Soc. Proc.,' A, vol. 86, p. 610 (1912); and A, vol. 88, p. 313 (1913).
(6) Devaux, 'Ann. Rep. Smithsonian Inst.,' 1913, p. 261.
(7) Labrouste, 'Ann. de Phys.,' vol. 14, p. 164 (1920).
(8) Clark, 'The Determination of Hydrogen Ion Concentration,' 1920.
(9) Harkins and others, 'Journ. Amer. Chem. Soc.,' vol. 39, pp. 354, 541 (1917).
(10) Marcelin, 'J. de Physique,' vol. 1, p. 19 (1914).
(11) Langmuir, 'Proc. Nat. Acad. Sci.,' vol. 3, p. 141 (1917).
(12) Langmuir, 'Journ. Amer. Chem. Soc.,' vol. 38, p. 2221 (1916).
(13) Le Bas, 'The Molecular Volumes of Liquid Chemical Compounds,' 1915.
(14) Clowes, 'J. Physical Chemistry,' vol. 20, p. 426 (1916).
(15) Hartridge and Peters, 'Proc. Physiol. Soc.,' October 16, 1920.
(16) McBain and others, 'Trans. Chem. Soc.,' 1919, p. 1279; 'Nature,' March 10, 1921, and other papers.

The Total Heat of Liquid Carbonic Acid.

By C. Frewen Jenkin, C.B.E., M.Inst.C.E., and D. N. Shorthose, M.A.

(Communicated by Sir Alfred Ewing, K.C.B., F.R.S. Received April 18, 1921.)

The measurements described in the following paper were made at the request of the Engineering Committee of the Food Investigation Board, who requested the author to measure the total heats of CO_2 at pressures and temperatures above the critical point. These total heats may be read from the I ϕ chart published in a former paper,* but the part of the chart concerned is based on an unverified assumption, viz., that the specific heat at constant vol., C_v, is constant and equal to 0·214, and it was considered advisable to check its accuracy by direct measurements.

The measurements of the total heats were made in the same manner as those described in the former paper; the carbonic acid being pumped round a circuit at a uniform measured rate and heated through various ranges of temperature in an electrical calorimeter. Much of the apparatus was the same as had been previously used, but an improved calorimeter was used and several modifications were required to enable the measurements to be made at the higher temperatures and pressures required.

The general arrangement of the apparatus is shown in fig. 1. P is a motor-driven compressing pump. C is a gas cushion to reduce the shock and keep the flow more nearly uniform; it consists of a steel pipe heated at the top in a hot-water vessel. D is a cooling coil to reduce the temperature of the CO_2 to a steady temperature before it enters the calorimeter. E is the calorimeter described below. F is a water-cooled condensing coil. G,G, are two flasks, one hung on a steelyard which turns with 1/100 lb. and rings a bell when the lever falls; the other is used simply as a container. During a test the valves on the flasks are set so that CO_2 flows out of the weighing flask, through the pump and calorimeter and into the container. As each 2 lb. of CO_2 leaves the weighing flask the steelyard lever falls and rings the bell; a 2 lb. weight is then hung on the flask thus raising the lever again. When the test is finished the valves are reversed and the CO_2 pumped back into the weighing flask. During the preliminary adjustment, before the actual test begins, the valves are set so that the CO_2 circulates through the weighing flask (which is full). H is a standard pressure gauge graduated to 10 lb. per square inch and capable of being read to 2 lb. per square inch. It is connected through a coil of small pipe in a hot-water vessel which acts as a cushion to steady the needle. The rate of flow of the liquid to the pump is controlled by the two valves V_1 and V_2 in series. The pressure gauge, J, indicates the pressure between the valves, and the pressure gauge, K, indicates the pressure on the suction side of the pump. The two valves enabled the rate of flow to be controlled much more accurately than it could with a single valve; during a test the valve V_2 was not touched and small adjustments were made on V_1 so as to keep the pressure, shown by gauge J, constant. The pressure in the calorimeter is controlled by the valve V_3 and read on the gauge, H.

---

* 'Phil. Trans.,' A, vol. 215, p. 368.

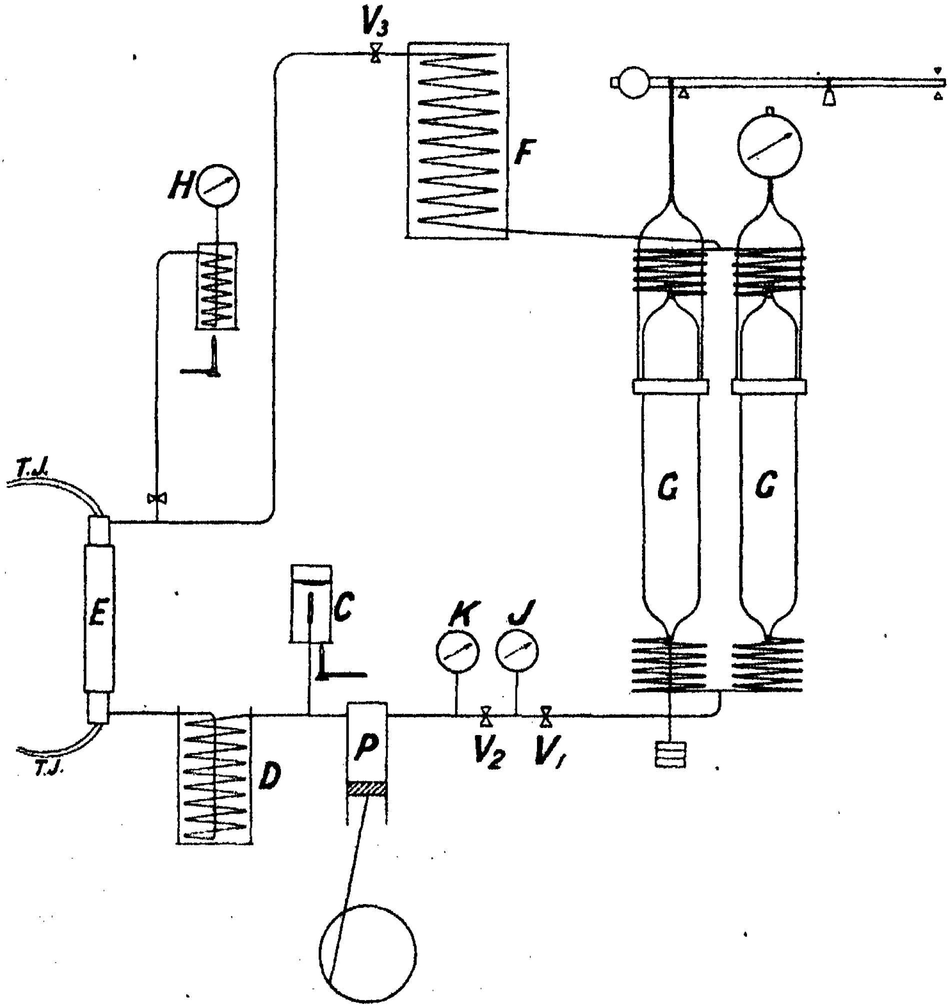


FIG. 1.

The cycle of operations on the $I\phi$ chart is shown in fig. 2. The condition of the liquid coming from the weighing flask is shown at "a." It expands through the throttle valves V_1 and V_2 to a pressure of about 200 lb. per square inch and is then in condition "b." It is then compressed adiabatically to the desired pressure, say 1,800 lb. as shown at "c." It is cooled in the coil, D, to about 12° C., to the point "d." It is then heated to the desired

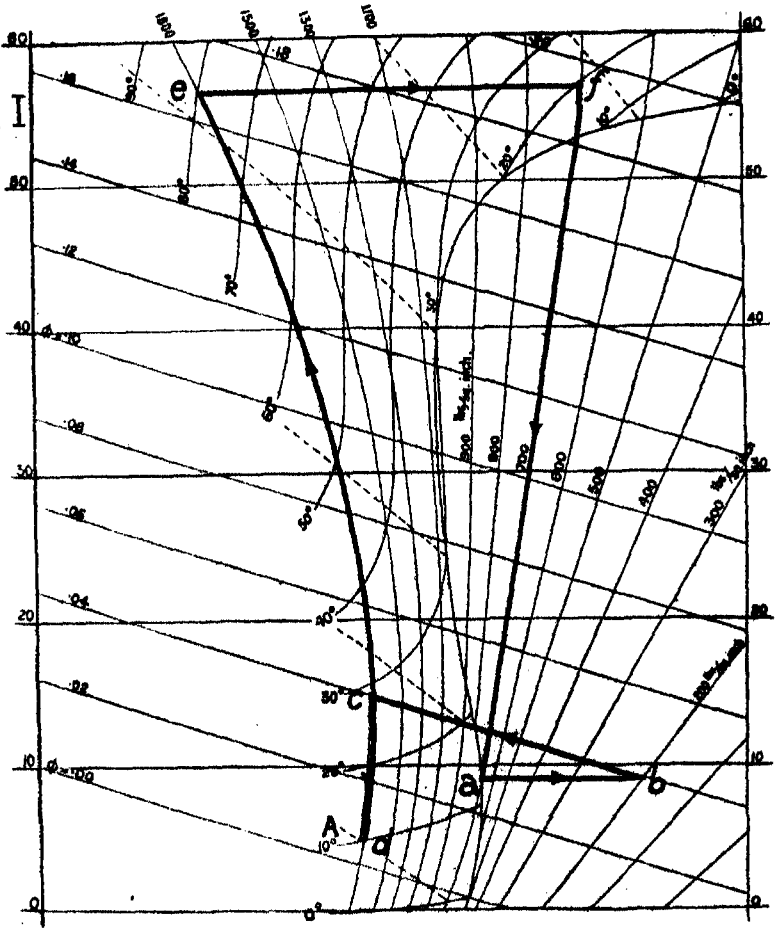


Fig. 2.—Part of I φ chart near the critical point.

temperature, to the point "e," in this diagram 80° C. It is then throttled through valve V_3 to the point "f," and cooled in the condensing coil, F, to the starting point "a." The total heat measured is shown on the chart by the vertical height between "d" and "e."

The calorimeter is shown in fig. 3. It consists of an outer steel tube, capable of standing the maximum pressure used, inside which is a bakelite

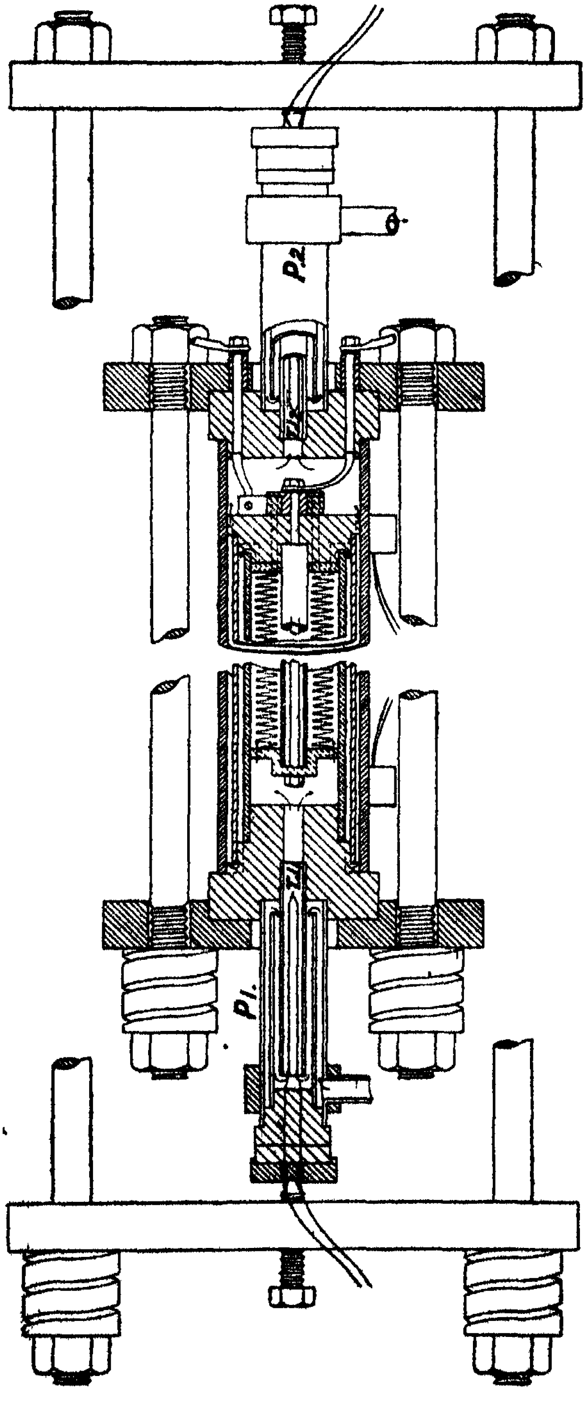


Fig 3 —Calorimeter

tube, and inside this again a fused silica tube. In the centre is a group of six nichrome steel-wire heating coils lashed round a small silica tube which supports them. The electrical connection to the one end of the heating coils passes through the central silica tube. The entering CO_2, which is at about atmospheric temperature, passes first along the outer space just inside the steel tube; it then returns along the space inside the bakelite tube, and finally passes down inside the silica tube, and is heated by the wire coils. By this arrangement the temperature of the outer steel tube is kept low, and radiation losses almost eliminated.

The CO_2 enters and leaves the calorimeter by two similarly arranged self-jacketed pipes, p_1 and p_2, enclosing the thermo-junctions, TJ_1 and TJ_2, which measure the temperatures before and after heating. The ends of the calorimeter were first made of hard vulcanite, but this softened at about 70° C., and vulcanite had to be abandoned for the hot end; a disk of bakelite was substituted, and answered admirably.

The thermo-junction wires were originally taken through the ends of the entering and outgoing pipes in the manner described in the former paper. A small button of brass was soldered on the wire; this was drawn through a thin rubber tube, which in turn was drawn into a conical hole in the vulcanite plug, as is shown in fig. 4. This arrangement was found to last only for a few hours, under the higher pressures, before the rubber tube failed. A new design, shown in fig. 5, was then tried; it consists of a small brass bolt with conical head through the centre of which the wire passes, being soldered into it. This formed a tight joint and appeared to

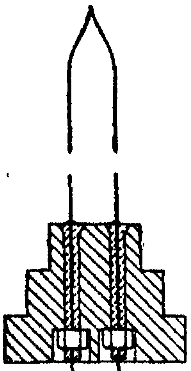


Fig. 5.

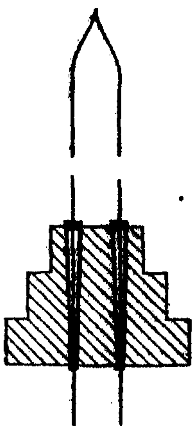


Fig. 4.

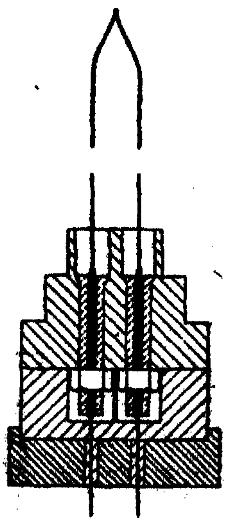


Fig. 6.

be satisfactory, but fortunately, a discrepancy in temperature readings, referred to below, called attention to a serious error which this bolt introduced. It will be noticed that the bolt virtually "short-circuits" a short length of the eureka wire; this has no effect unless there is a temperature gradient in the wire at that place, but if there is a temperature gradient it has the effect of reducing the temperature reading by approximately the difference of temperature between the two ends of the bolt. In other words, the use of the bolt is nearly equivalent to the introduction of an equal length of copper-wire in the eureka lead, which will give an opposing E.M.F., due to the difference in temperature between its two ends. This point was fully investigated by making a series of bolts of different lengths and soldering them on one of the wires of a thermo-junction a few inches apart and then calibrating the junction in an oil bath. No error was introduced except when the bolt was near the surface of the oil, *i.e.*, where there was a large temperature gradient in the wire, and then the error was roughly proportioned to the length of the bolt. The actual error will depend on many factors, but to show its importance the following may be quoted. The error was 3° C. for a bolt 20 mm. long in an oil bath at 40° C., the cold junction being about 10°, *i.e.*, the error was about 10 per cent. of the temperature difference.

The first method adopted for eliminating the error was to solder the bolt at one end only and to insulate the rest of its length from the wire. This, on trial, was found to be only partially successful, and further investigation showed that the thin layer of solder used to "tin" the wire was sufficient to cause a serious effect. Finally, the wire was coated with glue and blacklead to prevent the "tinning" from spreading, and a very short length cleaned and soldered to the bolt, the rest of the bolt being insulated from the wire by three layers of cigarette paper, shellac'd on (as shown in fig. 6). This proved perfectly effective. This small difficulty has been described at length because the presence of the errors may be so easily overlooked; it is purely a matter of chance whether it is found when the thermo-junction is calibrated.

The inlet and outlet temperatures were measured by the above-mentioned eureka-copper thermo-junctions and a potentiometer. The junctions were calibrated against a standard thermometer calibrated at the N.P.L., and a calibration curve plotted to a large scale.

The body of the calorimeter is surrounded when in use with a 2-inch cork jacket, given for this research by the Anglo Delta Slab Company. The temperature of the steel tube is measured by two Hoskin's alloy thermo-junctions, one at each end. A radiation test was made on the calorimeter, and showed that the loss of heat was 7·46 watts when the temperature difference between the steel tube and the air outside the cork jacket was 26° C. The maximum correction in any experiment for radiation was only 0·85 per cent. of the measured heat.

The electrical power supplied to the calorimeter was taken from the town mains (100 volt. D.C.) and regulated by adjustable resistances in series with the heating coils. It was measured by means of a Siemens precision wattmeter, having two ranges of current and three ranges of voltage. The wattmeter is divided into 150 divisions, and can be read to 1/10 division. The wattmeter was calibrated against standard resistances and cadmium cells which had been freshly calibrated by the N.P.L. Fluctuations in the town voltage gave considerable trouble; it was found necessary to watch the wattmeter continuously and make fine adjustments as required to keep the power constant.

Preliminary trials showed that it would not be possible to keep the conditions of a test steady without some indicator showing the upper temperature to which the CO_2 was being heated in the calorimeter. A Hoskin's alloy thermo-junction was therefore tied to the outlet pipe and connected to a millivoltmeter, but this was found not to be sensitive enough. To increase the sensitiveness the thermo-junction was connected to a d'Arsonval galvanometer, the "cold junction" being put into a well-jacketed tank of water which was heated to and kept at the temperature to which it was desired to heat the CO_2. This differential arrangement answered admirably. The galvanometer gives a deflection of 125 mm. per degree centigrade (the whole 500 mm. scale corresponding to 4° C.). A paper scale marked in half degrees in bold figures was used, which could be seen across the laboratory and enabled all the four observers to see any variations which occurred in the top temperature. It was a disagreement between this indicator and the temperature given by the potentiometer which called attention to the serious error (which has been already described) introduced by the bolts in the thermo-junction leads.

The tests were carried out in the following way. The jacketed tank containing the cold junction for the indicator was heated to the temperature at which the test was to be run. The pump was started and the pressures approximately adjusted. The current was switched on to the heating coils in the calorimeter, and adjusted so that the power was about that estimated to give the required rise of temperature when the rate of flow of CO_2 was 1 lb. per minute. The rate of flow was then adjusted by closing the valve V_2 till the temperature reached the desired point, as shown by the indicating galvanometer. After running for a short time, to allow all the conditions to get steady, the test was started.

During the test, readings were taken every minute of the following quantities: inlet temperature (potentiometer), outlet temperature (potentiometer), electrical power, outlet pressure.

The two jacket temperatures were read two or three times during a test. The times were noted when the bell rang, indicating the completion of each 2 lb. of CO_2. The total quantity of CO_2 passing through the calorimeter during a test was usually 20 lb.

The tests at the lower temperature were run at rates of about 1 lb. per minute, but at the higher temperature this rate had to be reduced, to prevent surging, shown by periodic variations of pressure and temperature. A slight increase of rate of flow has two effects. (1) A *lowering* of the final temperature, owing to the increased quantity to be heated. (2) A raising of the pressure, which reduces the value of I per pound, and consequently produces a *rise* of the final temperature. When the second effect is the larger the temperature rises, and the CO_2 expands, thus further increasing the pressure, and the condition is unstable. Considerable skill was required in manipulating the valves, to avoid periodic fluctuations due to this instability. At high rates the period of the fluctuations was too fast to control, but at lower rates it could be controlled satisfactorily.

Measurements were made at 900, 1000, 1100, 1200, 1300, 1500, and 1800 lb. per square inch pressure between temperatures of about 12° C. and 100° C. The results, after reduction, are shown by the dots in fig. 7. Smooth curves have been drawn through the points, and curves interpolated for 1400, 1600, and 1700 lb. per square inch. The values of the total heat for every 5° C. for each of the above pressures, taken from the curves, are given in the following Table. The observations at the higher temperatures are not quite as accurate as it was hoped to make them. The inaccuracy was due to the difficulty in keeping the top temperature steady. Repeated attempts were made to get steadier conditions by varying the rate of flow and by modifying the way in which the variations in the E.M.F. of the town supply were allowed for, but it was not found possible to get entirely satisfactory results.

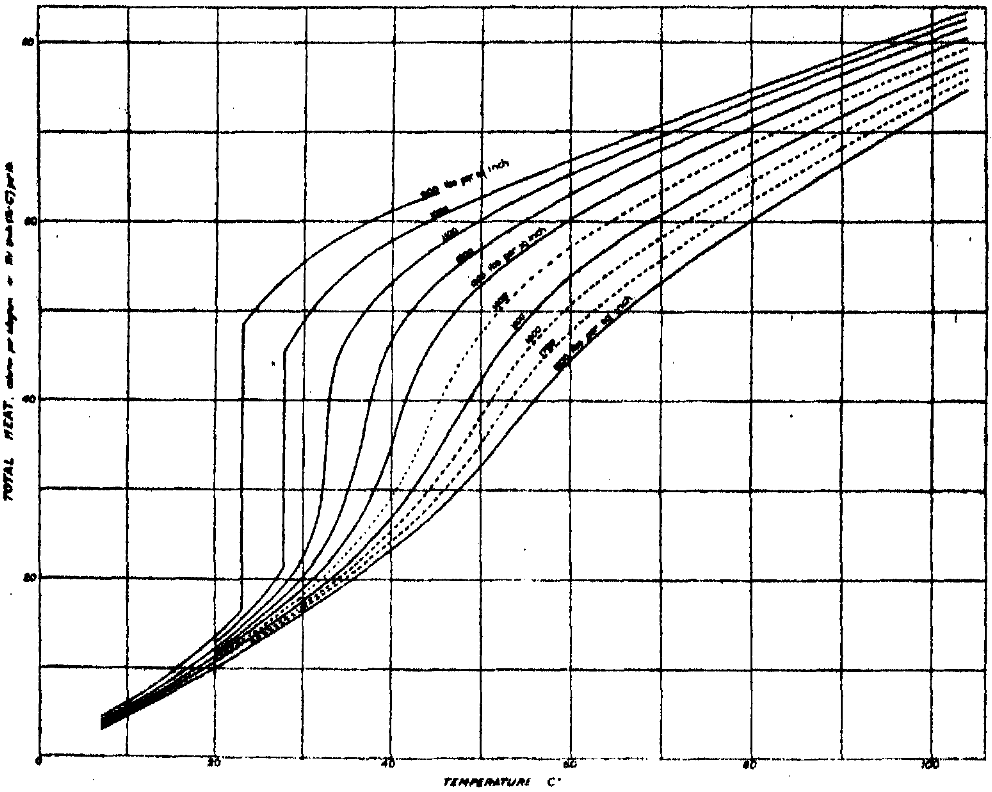


FIG. 7.

Total Heat I of CO_2 (Calories per kg. or lb. C.° per lb.).

Temp. ° C.	Pressure lb. per sq. in.									
	900	1000	1100	1200	1300	1400	1500	1600	1700	1800
0	0·34	0·23	0·14	0·06	—	—	—	—	—	—
5	3·21	3·05	2·90	2·77	2·64	2·55	2·47	2·38	2·29	2·19
10	6·16	5·92	5·70	5·51	5·34	5·19	5·02	4·89	4·75	4·60
15	9·5	9·1	8·8	8·5	8·2	7·9	7·7	7·5	7·3	7·2
20	13·5	12·9	12·3	11·8	11·4	11·0	10·7	10·4	10·2	10·0
25	50·8	17·5	16·4	15·6	15·0	14·4	14·0	13·6	13·4	13·1
30	55·4	48·9	22·4	20·2	18·9	18·3	17·5	17·1	16·7	16·2
35	58·6	54·4	47·3	29·7	24·6	22·9	21·6	20·8	20·2	19·7
40	61·1	57·8	53·2	46·9	35·1	29·2	26·8	25·4	24·2	23·3
45	63·1	60·4	57·1	53·1	47·2	39·2	34·1	31·0	28·9	27·4
50	—	62·8	60·2	56·9	52·9	47·6	42·2	38·3	35·2	32·7
55	—	64·7	62·9	60·1	56·8	53·1	48·9	45·4	42·3	38·9
60	—	66·8	65·2	63·0	60·2	57·1	53·9	50·8	47·7	44·5
65	—	68·8	67·4	65·6	63·1	60·3	57·6	54·8	52·0	49·3
70	—	70·8	69·5	68·0	65·7	63·3	60·8	58·2	55·7	53·2
75	—	72·7	71·5	70·2	68·2	66·1	63·8	61·5	59·1	56·8
80	—	74·6	73·6	72·3	70·6	68·7	66·6	64·4	62·3	60·1
85	—	76·5	75·6	74·4	72·8	71·0	69·3	67·3	65·3	63·3
90	—	78·4	77·5	76·4	74·9	73·4	71·8	70·0	68·2	66·5
95	—	80·2	79·4	78·3	77·0	75·6	74·1	72·6	71·0	69·5
100	—	82·0	81·2	80·2	79·0	77·7	76·4	75·1	73·7	72·4

Comparing the results with the Iϕ chart, published in the former paper (*loc. cit.*), it will be found that the chart gives slightly too low results, the largest error being in the left-hand top corner. The following examples will show the extent of the differences:—

Pressures.	100° C.	70° C.	40° C.
	per cent.	per cent.	per cent.
1800	7·0	6·7	4·2
1100	4·6	3·4	4·1
900	—	—	2·5

The results show that both the pressure lines and temperature lines require to be slightly shifted in the chart.

The results of the present investigation, combined with those in the author's last report,* are theoretically sufficient for the evaluation of the specific heat at constant volume, C_v, which has hitherto been assumed to be constant, as stated at the beginning of this paper. It is not easy to deduce from the data very accurate values for C_v, but the calculations appear to show that they are far from constant. This point is being further investigated, as it appears to be of considerable theoretical interest.

---

* 'Roy. Soc. Proc.,' A, vol. 98 (1920).

The Energy Involved in the Electric Change in Muscle and Nerve.

By A. V. Hill, F.R.S.

(Received April 8, 1921.)

[This paper is published in 'Proceedings,' Series B, vol. 92, p. 178 (No. B 645).]

A Quantum Theory of Colour Vision.

By J. Joly, Sc.D., F.R.S.

(Received April 8, 1921.)

[This paper is published in 'Proceedings,' Series B, vol. 92, p. 219 (No. B 646).]

The Time Interval between Absorption and Emission of Light in Fluorescence.

By R. W. Wood, For. Mem. R.S., Johns Hopkins University, Baltimore.

(Received June 12, 1921.)

[Plate 5.]

The more recent theories of the absorption and emission of light, as in cases of resonance radiation and fluorescence, make the two processes quite distinct. According to Bohr the absorption of radiation of frequency corresponding to D_2 by non-luminous sodium vapour, is associated with the passage of an electron from an inner to an outer stable orbit; resonance radiation resulting when the electron returns to the original orbit. The time interval between these two processes is probably too short to be detected experimentally, for the faintness of the light makes it appear doubtful whether methods, similar to the one devised by Abraham and Lemoine, can be used in these cases. In the case of the fluorescence, or rather phosphorescence, of mercury vapour, I have succeeded in measuring the time interval, the vapour remaining non-luminous during the process of the absorption of light, and bursting into luminosity about 1/15000th of a second later. This, I believe, is the first case ever observed of a photo-luminescent body remaining dark during the period of excitation.

In the study of processes of this nature two methods are available: we may give the substance a very high velocity of translation or rotation and cause it to pass through a very narrow and intense beam of light, or we may keep the substance stationary, illuminating it with flashes of very brief duration and examining it during the moment of excitation, or at definite moments later, through a perforated disc rotating at high speed (phosphoroscope). Both methods have been employed in the present work, the former being preferable in some respects.

The fluorescence of mercury vapour is bluish green in colour and results from the absorption of ultra-violet light in the region of very short wave-lengths (1850–2100). It would seem probable, in view of Steubing's results on the ionization of the fluorescing vapour, that the absorption of light causes the expulsion of an electron, the fluorescence resulting from its return to the atom. The phenomenon, however, is probably not as simple as this, for monatomic mercury vapour in equilibrium with the metal, or when no fluid metal is present, does not exhibit fluorescence at any density or temperature.

The fluorescence appears only in the case of vapour *freshly liberated from the fluid metal*, probably diatomic mercury. This is indicated by the circumstance that when an exhausted quartz bulb containing a drop of mercury is very gradually heated in an electric oven, fluorescence is barely visible at 170° and rises to great intensity at 180° (certainly a ten-fold increase in intensity). The vapour density, with this rise of temperature, increases from 8 mm. to 11 mm. If, now, the oven is allowed to cool, fluorescence disappears as soon as the temperature has fallen two degrees, *i.e.*, as soon as condensation begins.

The fact has been still more clearly shown in some recent work done in collaboration with Mr. J. S. van der Lingen, in which the mercury is contained in a double bulb connected by a short tube. One bulb is at 200°, the other at 205°, the mercury gathering in the cooler bulb. Both are illuminated by a spark and no fluorescence is observed. If the hot bulb is now suddenly cooled by an air blast, the other bulb immediately shows brilliant fluorescence, for the mercury in it immediately begins to vapourize. The absorption, however, appears to be a characteristic of the atom, at least I have never been able to detect any difference between the absorption of freshly forming vapour, and that of vapour in a static condition.

The time interval which elapses between the absorption and emission of the light can be easily observed in the following way: a tube of fused quartz of the form shown in Plate 5, fig. 1 (with a total length of 15 cm. and a bore of 4 mm.) containing about 1 c.c. of mercury, is exhausted and sealed. The tube should be heated during the exhaustion and the metal boiled. It is mounted in a clamp-stand and the mercury bulb heated by a Bunsen flame about 1 cm. in height. An aluminium spark (condensed discharge) is arranged as shown in the figure, the light passing through a hole 2 mm. in diameter in a thin metal plate mounted close to the tube. The cylindrical beam of light should pass through the centre of the tube, which can be accomplished by observing the circular illuminated patches where it passes through the walls. On lighting the burner, as soon as active distillation of the mercury commences, we observe a faint greenish light filling the entire tube. This is the phenomenon observed by Phillips* with a tube of similar form, excited, however, by the 2536 line of the mercury arc.

As the temperature of the "condenser" end of the tube rises, the velocity of the vapour becomes less, the luminosity draws down towards the illuminated region, and presently appears in the form of a beautiful green flame, concave on the under side as shown in Plate 5, fig. 1. The flame form is due obviously to the high velocity of the vapour along the axis of the tube, and the low velocity close to the wall. Absolutely no sign of luminosity is seen in the region traversed by the exciting beam, except a trace on the inner wall of the tube where the velocity of the vapour is very low. At the centre of the tube the dark region extends 2 or 3 mm. above the beam. The form of the flame gives us a means of measuring the variation in the velocity of the vapour across a diameter of the tube. The "flame" appears at its best only when the Bunsen burner is at exactly the right height; if it is supplying heat too fast the fluorescence draws down into a flatter patch and may even enter the illuminated region. This results from too great back-pressure, which diminishes the velocity of the vapour.

---

* Phillips, 'Roy. Soc. Proc.,' A, vol. 89, p. 39.

A photograph of the fluorescent "flame" is reproduced on Plate 5, fig. 4. This was taken with a glass lens, which excludes the powerful ultra-violet radiation (2536), which is also radiated by the "flame." The patch of light at the left is caused by the reflection and scattering of the exciting beam by the wall of the quartz tube, the fainter patch to the right marking the exit of the beam. Special precautions are necessary in the study of the phenomenon, with respect to this ultra-violet, monochromatic radiation. The "flame" must be photographed with a quartz lens and an objective prism of quartz, and the luminous patches of spark light scattered by the tube carefully screened off. Unless an achromatic quartz fluorite objective is used, the photographic plate must be inclined if the images corresponding to different wave-lengths are all to be in focus. This broadens the images of course. A photograph obtained in this way is reproduced in Plate 5, fig. 5. The "flame" evidently has the same general form in the monochromatic 2536 radiation, as in visible light. The relatively great intensity of this radiation is clearly brought out in the photograph.

If the spark is placed close to the bulb and the screen removed, the entire bulb appears filled with green fluorescent light, but if the eye is brought to the level of the mercury surface, a dark zone is seen between the free surface of the mercury and the luminous vapour, which reminds one strongly of the Crookes dark space in vacuum tubes. The width of this dark zone is about 1 mm. This is what we should expect, for the velocity with which the vapour leaves the mercury surface is only about one-third of the maximum velocity in the tube, on account of the larger cross-section of the vapour column in the bulb. If we had any means of measuring the velocity of the mercury vapour, we could at once determine the time intervals involved, but this is not easy. A rough approximation was made by calculating the volume of metallic mercury which distilled over in a given time (about 0·75 c.c. in six minutes) and determining the density by photographing the 2536 absorption line, comparing the width of the line with the widths shown in photographs made with a sealed tube at various constant temperatures. This was necessary as we cannot assume that the density of the vapour corresponds to the temperature at which the hot end of the distillation tube is held, for it depends upon the rate at which heat is actually supplied to the fluid mercury in the bulb. Estimates made in this way indicated that the average velocity was of the order of 10 metres per second. The velocity along the axis of the tube might easily be double this.

Owing to the uncertainty of these calculations, it seemed desirable to measure the time intervals directly, and a special type of high-speed phosphoroscope was designed for the further study of the phenomenon.

Some experiments were also made with an exhausted spherical bulb of quartz, 4 cm. in diameter, containing a few drops of mercury, and rotated at very high velocity. The whirling bulb was heated by a small blast lamp, and the spark focussed on the wall by a quartz lens. The fluorescent spot was not, however, displaced by the rotation, so far as could be seen. The reason for this will appear later, as it has been found that with great vapour density in bulbs the time interval between absorption and emission is much briefer than in the case of the inverted U-tube with distillation going on.

The resonance radiation excited by the monochromatic radiation of wave-length 2536 was next investigated. This is shown by the vapour of mercury at room-temperature, and probably does not depend upon distillation. This point has not, however, been proven as yet, for evaporation and condensation are considered as going on all the time at the surface of a liquid in contact with its saturated vapour. Resonance radiation should be looked for in a bulb containing no liquid mercury.

The first experiments were made with the inverted U-tube, the bulb being heated by a small vessel of hot water. No trace of any displacement of the radiating mass of vapour could be detected. A vertical tube, 2 cm. in diameter, and 10 cm. long, with flat ends (very highly exhausted) was then used, the upper end being provided with a cup of paper packed with cotton soaked with liquid air. The mercury at the bottom was at room-temperature. The cone of exciting radiation was obtained with a quartz monochromator and water-cooled mercury arc. The cone of resonance radiation was found in coincidence with the exciting cone, showing that the time interval between absorption and emission in this case was too small to be detected by this method.

We will now take up the experiments with the phosphoroscope, which yielded much more definite results.

The Synchrono-Phosphoroscope.

In order to determine with greater precision the interval of time elapsing between the absorption of the light and the emission, a spark phosphoroscope of special design was constructed. With this instrument I have been able to detect phosphorescence of a duration of less than 1/400,000 second, which is about ten times as brief as any observed phosphorescence of which I can find record.

A thin disc of Bakelite (similar to hard rubber), 45 cm. in diameter, was mounted on the shaft of a synchronous motor, operated by the same 60 cycle alternating current which fed the primary of a small 30,000 volt transformer. A small strip of thin sheet brass, 3 cm. long, bent out into two points, was screwed to the disc (parallel to a radius) near the rim. Two pointed brass wires, wound in short spirals and mounted on a block of hard rubber, were brought up close to the points on the disc, and the circuit arranged as shown in Plate 5, fig. 2. K is a capacity of two small Leyden jars, A the illuminating spark. A few trials are necessary to find the proper place at which to mount the pointed rods, in order to complete the circuit at the moment when the potential of the transformer is near a maximum. By carefully adjusting the distance between the fixed and rotating points, it is possible to fix the timing of the sparks so accurately that no motion of the little brass strip (illuminated by the light of the sparks which pass between its points and the wires) can be detected. The speed of the disc was 1800 revolutions per minute, yet it appeared absolutely stationary in a dark room. If the adjustment of the gaps is not right, the brass strip appears to dance up and down, in an irregular manner, with an amplitude of perhaps a millimetre, and the satisfactory operation of the instrument, at least for the study of the briefest time intervals, depends upon the complete elimination of this irregularity. The elasticity of the spirals obviates the danger of breaking the points by actual contact, which occurs in making the adjustment. A radial slit, 1 mm. in width, was cut in the disc near the rim, and a small quartz bulb containing a droplet of mercury, or the U-tube previously described, was mounted behind the disc and viewed through the slit, as the disc revolved. The illuminating spark (between aluminium rods) was mounted close to the quartz bulb.

It was found that, if the eye was brought up close to the disc at such a point that the sparks flashed at the moment when the moving slit crossed the pupil of the eye, the outline of the bulb or tube, illuminated by the spark light, could be seen, but no trace of the green phosphorescence of the vapour could be detected. This came into view when the eye was moved a trifle in the direction in which the slit was moving, the bulb disappearing, of course. It was found that the distance between these two positions was so small that the diameter of the pupil of the eye gave trouble, and a second slit was mounted close to the disc, and arranged so that it could be moved to the right or left by small increments. When everything was in the best adjustment, it was found that a movement of the second slit through a distance of about 2 mm. brought the phosphorescence into view, in the case of the inverted U-tube, while, with the bulb, the distance was so small that some uncertainty was felt as to whether the effect was really present. The reason for this will appear presently. This means that the time interval between the flash of the spark and the emission of the fluorescent light is of the order of 1/20,000 second.

The operation of the phosphoroscope in this form was not quite satisfactory, owing to the fact that very narrow slits had to be used for the measurement of the shortest time intervals, a disadvantageous circumstance in consideration of the comparative faintness of the light sources employed. A great improvement in the performance of the instrument was obtained by cementing a small lens (1 cm. in diameter and 3 cm. focus) in a hole in the disc near the rim. This lens formed a real image of the source of light, which was viewed with a short focus lens or eyepiece, the image, of course, moving with the velocity of the lens. The lens is inserted in such a position on the disc as to ensure its moving in a nearly horizontal direction at the moment the spark occurs, *i.e.*, the lens should be at the top of the disc when in the sparking position. It was thought that this device would show the motion of the fluorescent vapour by giving an oblique image of a vertical slit, backed by the fluorescent "flame." Owing to the inconvenience of operating the mercury tube close to the disc, it was placed at a distance of about a metre, and an image of it projected at a point about 10 cm. in front of the small lens, with the disc set in the sparking position. This arrangement gave most satisfactory results. The vertical image of the slit was plainly visible by the spark light scattered by the walls of the quartz tube, while the green oblique image appeared completely detached from it, and inclined at an angle depending upon the velocity of the vapour, as shown on Plate 5, fig. 6: We have here complete data for determining the period of darkness which precedes fluorescence, by two different methods:—

First: We measure the distance between the base of the vertical image of the slit and lower end of the oblique band. This was found to be 2·5 mm., and, since the image moves 1 mm. in 1/39,000 second, the time interval between the excitation and the emission of the first fluorescent light is 1/15,600 second.

Second: By measuring the obliquity of the displaced image we can calculate the velocity of the vapour. The flash of light from the spark passing through a small hole gives an instantaneous excitation to a small mass of mercury vapour which is moving upward with a high velocity. By the time it bursts into luminosity it has already moved up a short distance, and as it continues to ascend, the moving lens carries its image to the right. The tangent of the angle which the oblique image thus formed makes with a horizontal line is the ratio of the velocity of the vapour to that of the lens, and this angle was found to be about 23°. This gave 17 metres per second as the velocity of the vapour. Observations of the fluorescent flame showed that the luminosity commenced about 1·5 mm. above the excited region. This gives us for the time interval 1/11,400 second, a value somewhat smaller than the one obtained by the first method. The former is probably more nearly exact. A small bulb was now substituted for the inverted U-tube, heated by a column of hot air rising from an asbestos chimney. The spark was placed close to the bulb, and the fluorescence occurs chiefly on the inner surface on the side facing the spark (superficial fluorescence) owing to the density of the vapour.

The whole slit was simultaneously illuminated by the fluorescent light under these conditions, and it was expected that the image seen with the moving lens would appear as in Plate 5, fig. 6. It was found, however, that the band of green fluorescent light was joined to the image of the slit formed by the light (white) of the spark scattered by the walls of the bulb. This explained at once the difficulty found in getting results by the first method, in which the bulb was viewed through a narrow slit in the disc. The greater density of the vapour in this case evidently reduces the time interval between absorption and emission to something less than 1/40,000 of a second, as this would give a displacement of a millimetre, which could be observed without difficulty. A vertical tube of quartz about 2 cm. in diameter and 10 cm. long, was now substituted for the bulb, and the pool of mercury at the bottom heated by a small flame. The spark was placed close to the front wall of the tube. The phosphoroscope now showed a beautifully displaced band, as shown on Plate 5, fig. 7. The time interval in this case was about the same as in the other experiment, 1/15,000 second.

Experiments with Phosphorescent Solids.

The phosphorescence of solids can obviously be investigated by the same method, but a simpler arrangement was found preferable, namely, to make the solid itself the moving source. Uranium glass was examined first. A thread of the glass measuring about 0·1 mm. in diameter was drawn out and a small fragment cemented to the edge of the disc, about 5 mm. projecting beyond the rim. The spark was placed as close as possible to the glass thread, behind a screen of black paper, which kept the direct light out of the eye. On setting the disc in rotation, and examining the thread of uranium glass with a short focus lens, it was found that the broad green phosphorescent band joined the white image of the thread (formed by refracted and reflected light). I feel sure that a dark band between the two 0·1 mm. wide could have been detected if it existed. This proves that the time interval in this case must be less than 1/400,000 second.

Barium platino-cyanide was next investigated; a long crystal, 0·5 mm. wide, showed no trace of phosphorescence, the crystal appearing quite sharp in its brilliant green fluorescent light; a needle-shaped crystal, 5 mm. long and 0·1 mm. in diameter, was then cemented to a small strip of thin lantern slide glass, with collodion, and the glass plate attached to the rim of the disc with sealing wax, as shown in Plate 5, fig. 3. This gave the very surprising result that the total duration of the phosphorescence was only about 1/400,000 second. The crystal appeared accompanied by a phosphorescent band of a width almost exactly equal to that of the crystal (0·1 mm.). This came pretty close to being fluorescence.

A number of other substances were examined, and the results are given in the following Table :—

Substance.	Width of phosphorescent band.
Uranium glass	One-quarter of revolution of disc.
Willemite	*Three-quarters of revolution of disc.*
Balmain paint	Blue-green phosphorescence lasting 1/40,000 second. Violet phosphorescence of long duration too faint to appear.
Anthracene	Small crystals cemented to glass, ¼ mm. (1/160,000 second).
Barium platino-cyanide	0·1 mm. (1/400,000 second).

Fluorescent Solutions.

In the case of liquids which are available in sufficient bulk, another method is preferable, for we can give a liquid jet a much higher velocity than anything that can be obtained with a phosphoroscope. I made a few experiments with solutions of fluorescene, using jets of comparatively low velocity (having no facilities for high pressure), and focussing sunlight on them. No displacement of the fluorescent spot could be detected.

Some experiments were then made at the University of Wisconsin, in collaboration with Prof. C. E. Mendenhall, during my visit to Madison in December. We used a high pressure, six-cylinder pump, and obtained a jet velocity of about 200 metres per second, with a fine glass nozzle about

0·2 mm. in diameter. More recently, Prof. Mendenhall has increased the velocity to 230 metres per second, and, by blackening one side of the jet tube, leaving a small clear space for the entrance of the sunlight, has assured himself that there is no displacement as great as 0·1 mm. (observing the fluorescent patch with a short-focus lens). This means that the duration of the fluorescence is less than 1/2,300,000 second.

In the case of the fluorescence of solutions, however, we must bear in mind that the emission of light is accompanied by, if not the direct result of, the destruction of the fluorescent body. This has been known for some time in the case of certain substances (anthracene, for example), and its general application has recently been pointed out by Perrin, who showed that thin films of fluorescent solutions were bleached by the action of light, and lost their power of exhibiting fluorescence.

I have examined this question on a somewhat larger scale, concentrating sunlight with a large lens, and have found that fluorescene, eosine, rhodamine and other similar bodies are transformed into non-fluorescent coloured solutions, with absorption bands of different form from the original bands, and located in different regions of the spectrum. If substances of similar optical properties can be prepared by chemical or electrical means (oxidation or reduction, for example) from the original substances, we shall be able to tell, with tolerable certainty, just what change results from the action of the light. It seems a remarkable hypothesis, at all events, that the emission of light in these cases results from a molecular explosion produced by the radiation, rather than from the return of an expelled electron. If this is the case, the brevity of the phenomenon is accounted for.

There seems to be one rather curious circumstance, to which attention has never been drawn so far as I am aware, namely, that all fluorescent solutions are solutions of organic compounds. The only exception to this rule that I know of is in the case of very concentrated solutions of some of the uranyl salts, which show a very feeble greenish fluorescence when illuminated in a dark room with a powerful beam of concentrated sunlight passed first through dense cobalt glass.

The feebleness of the fluorescent light, and the great concentration of the solution, suggests that only a very small percentage of the molecules are engaged in the process, and the idea occurs to one as to whether the phenomenon may not perhaps be associated with the natural breakdown of uranium. These matters will be more fully discussed in another paper.

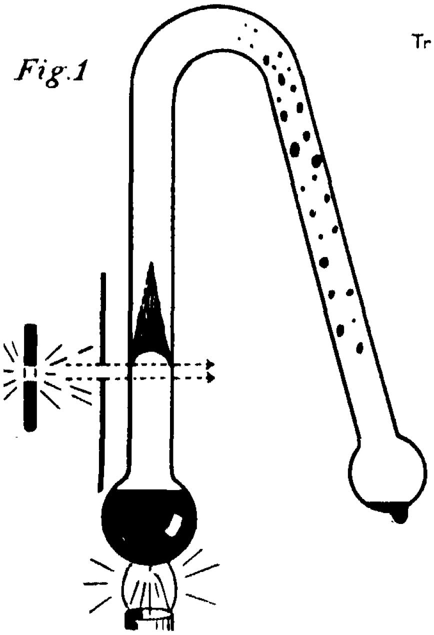


Fig. 1

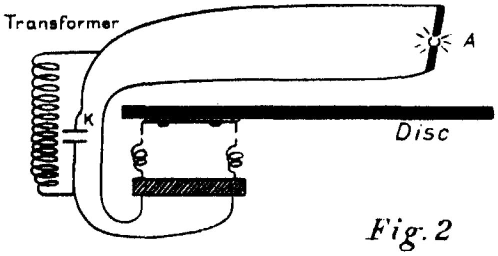


Fig. 2

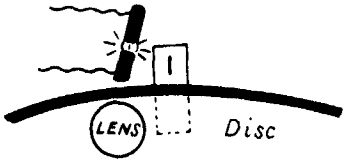
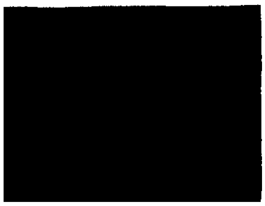


Fig. 3

Fig. 6



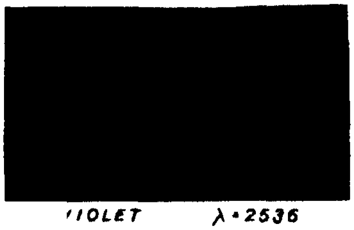


Fig. 5

Fig. 4

Fig. 7

Application of the Method of Abraham and Lemoine to the Study of Phosphorescence.

This method (see my 'Physical Optics,' p. 549) enables us to deal with time intervals of the order of 1/400,000,000 second. It depends upon the double refraction of bisulphide of carbon or nitro-benzol in an electrostatic field. Two plates are immersed in the liquid, and these form the plates of a condenser, which discharges at a spark gap placed behind a nicol prism and viewed through a second nicol, placed at extinction, through the space between the plates of the fluid condenser.

If the spark is close to the condenser, the light reaches it before the double refraction has disappeared, and is therefore passed by the second nicol. If, however, the light is made to traverse a path of about 2 metres, by reflecting it from a mirror, the double refraction has disappeared during the time taken by the light to traverse this distance, and the image is no longer visible.

With bisulphide of carbon the restored light is so feeble that it was obvious that the light of a crystal of barium platino-cyanide, placed close to the spark, would not be seen.

With nitrobenzol, in which the Kerr effect is sixty times as powerful, however, there seemed to be some hope of getting results. It was found that even freshly distilled pure nitrobenzol conducted too well to form a condenser. Phosphorus pentoxide changed it to a deep red liquid in a few hours, and lime did not prove effective in removing the traces of water which were assumed to give the trouble. Dehydrated copper sulphate, poured while still red-hot into the liquid, and shaken up with it at intervals for a few days, brought it into excellent condition. With this I had no difficulty in seeing the greenish image of a crystal of barium platino-cyanide placed very close to the spark. If it had remained dark for a period of say 1/10,000,000 of a second, its light would not have been passed by the second nicol.

There is room for further work along these lines. So far as I am able to say now, it appears that the duration of the double refraction with nitrobenzol is considerably greater than in the case of carbon bisulphide, for I could not get the spark to disappear even with a path of 15 metres.

I have been aided in this investigation by a grant from the Rumford Fund of the American Academy of Arts and Sciences.

A Study of the Glow of Phosphorus.—Periodic Luminosity and Action of Inhibiting Substances.

By Lord Rayleigh, F.R.S.

(Received June 17, 1921.)

§ 1. *Introduction.*

Many experimenters who have used cold phosphorus for the absorption of oxygen from air must have noticed the appearance of flickering clouds of luminosity when the action is nearly complete. These occur in the gas space and obviously represent a delayed action between the slight oxygen residue and the vapour of phosphorus.* This action has been discussed by Joubert,† who recognised it as due to the propagation of combustion in an explosive mixture. He obtained it in a more striking form by the slow leakage of air into an exhausted vessel containing phosphorus: his conclusion was that below a certain pressure, too small to measure, phosphorus vapour would not combine with oxygen. As the mixture of phosphorus vapour and oxygen became richer in oxygen by the inflow of the latter, the point was reached when combustion became possible and an explosion was propagated.

From the point of view of kinetic theory it seemed very strange that phosphorus vapour should behave thus. Joubert's view of the facts would seem to imply that the reaction of a molecule of phosphorus with oxygen was not dependant solely on the character and energy of molecular collisions, but also on the absolute value of the interval of time between them. These theoretical difficulties, and also the fascination of the experiments themselves, led me to attempt a further investigation.

§ 2. *Travelling Pulses of Luminosity. Need for the Presence of Water.*

The earlier attempts were directed to getting the effect in as striking a form as possible. The first experiment which gave results worthy of record was as follows: A horizontal glass tube 48 cm. long and 2 cm. diameter (fig. 1) was closed at the ends. Along the top side, at intervals of about 4 cm. a number of lateral openings were provided, to each of which was attached a long vertical capillary tube about 1 mm. in diameter, and 1 metre long. These tubes were open at the far end to the air of the room. Along the bottom of the main horizontal tube was a layer of phosphorus. This had been melted into position under a layer of water. After cooling the water was poured off as far as possible. The lateral tubes were then attached as described. The phosphorus at first presented a uniform luminous surface where it glowed in contact with the contained air, but after a time when the oxygen began to approach exhaustion, a flickering appearance set in. This developed, as the action proceeded, into a very definite and distinctive form. At the side entrances air was diffusing into the nitrogen atmosphere, which contained phosphorus vapour. Luminous pulses were observed to spring into existence at these side entrances, to divide, and to travel in opposite directions along the tube. The pulses were of the same diameter as the tube (2 cm.) and in thickness less than 1 cm. Their rate of travel was estimated at between 1 and 2 cm. per second. When two pulses originating at different places met one another, both were destroyed. It is easy to see why this should be the case. Suppose for distinctness of statement that the phosphorus vapour is in excess. Then each pulse as it passes leaves the space behind it exhausted of oxygen—its passage may be compared to the passage of an army living on the country, or a flight of locusts—no supplies are left in the rear. Another pulse meeting the first finds no oxygen to support its further advance or existence.

---

* In this paper I do not enter upon the question of whether the glow is due directly to oxidation of phosphorus, or of phosphorus trioxide. When, for brevity, the oxidation of phosphorus is referred to, this reservation should be understood.

† 'École Norm. Annales,' vol. 3, 1874, p. 209.

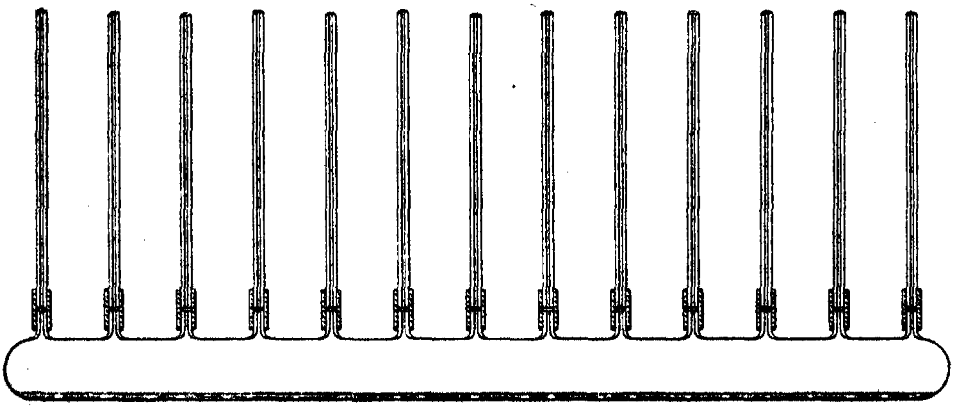


FIG. 1.

The arrangement of side entrances described was designed to provide, as far as possible, for a uniform admixture of phosphorus vapour and oxygen along the tube. On occasion a single pulse has been observed to traverse nearly the entire length. As a rule the pulses originated chiefly at two or three particular side junctions, perhaps owing to a favourable condition of the phosphorus surface for evaporation near these entrances; but the action was somewhat variable from day to day. A rise of the barometric pressure *would cause oxygen to enter more rapidly than by diffusion, and a change of temperature also had some effect.*

This apparatus continued to work well for a week or more, and was expected to do so almost indefinitely. But after a time it was noticed that the pulses were fewer and less brilliant, and that there were patches of stationary luminosity in the tube. After various possible explanations of the change had been examined, it was found to be due to the gradual drying of the phosphorus under the desiccating influence of the phosphorus pentoxide produced. On adding a few drops of water, the original condition was recovered.

For further examination of the effect of moisture, a tube was prepared, as in fig. 2, with a single side entrance, and a layer of phosphorus on the bottom. It was dried out by exhaustion on a mercury pump, with which a phosphorus pentoxide drying tube was connected. It was then filled up with commercial nitrogen (to save time), and the capillary entrance tube attached. There was now no trace whatever of the luminous travelling pulses. The glow, when air was entering, was perfectly steady, and took the form shown in fig. 2, *a*, *b*, *c*, which represent successive appearances as the air influx was gradually diminished. Any one of these stages could be maintained indefinitely.

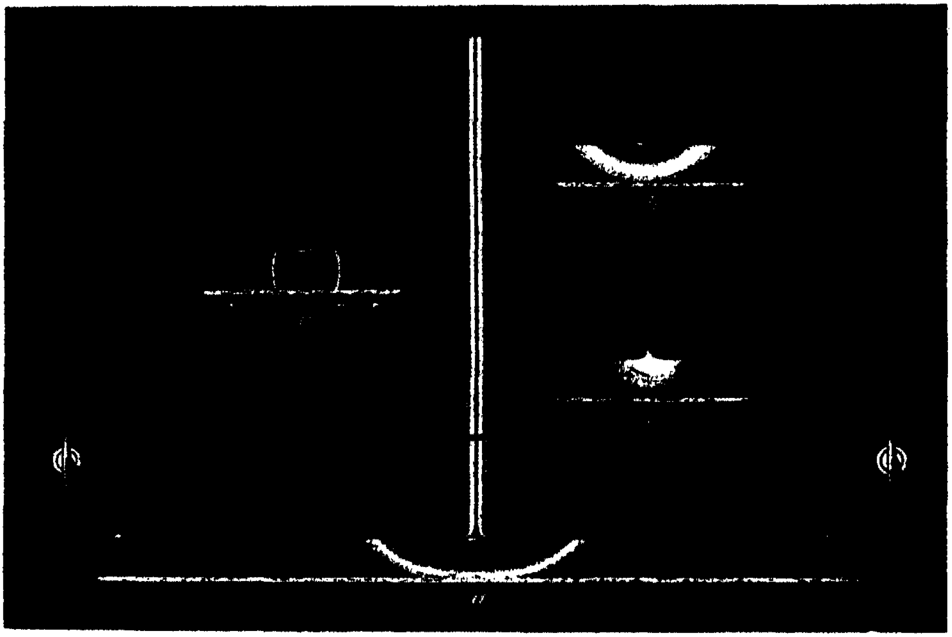


Fig. 2.

On adding a few drops of water, a constant succession of travelling pulses was substituted for the steady luminosity (fig. 2, *d*). In this case a pulse, starting from where the oxygen was most concentrated, divided, and travelled in each direction into a region where the oxygen became progressively scarcer. Thus the pulse gradually lost in intensity, and became extinguished after travelling a distance of perhaps 6 cm.

It is quite clear, then, that in these experiments the presence of water vapour is an essential condition for the periodic luminosity.

The quantity of water vapour necessary was next studied. In this case a tall narrow exhausted bottle was used, near the bottom of which was hung up a small piece of phosphorus—a slice cut from one of the ordinary commercial sticks. Air could be slowly admitted through a capillary tube, so as to raise the pressure a few millimetres per hour (fig. 3).

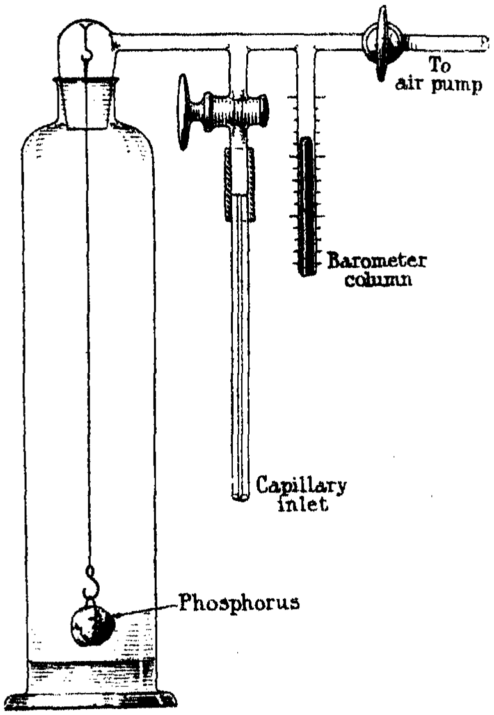


FIG. 3.

When the bottle contained a little water, intermittent luminous pulses travelled down it. Substituting strong sulphuric acid, the light at the top of the bottle, where air was diffusing into the phosphorus vapour, became quite steady. With acid of density 1·52 (vapour pressure 3 mm.), the flickering was much less conspicuous than with water only. Finally, acid of density

1·66 (vapour pressure 0·8 mm.) was about the densest that would leave any perceptible unsteadiness in the light.

The saturation pressure of phosphorus vapour at ordinary temperatures is 0·025 mm. To get the intermittent luminosity will require a pressure of water vapour about 100 times as great.

It will be seen, then, that the quantity of water vapour required is of quite a different order of magnitude from that concerned in Baker's experiments.[*] He found that prolonged drying over phosphorus pentoxide would stop the action of oxygen on phosphorus altogether.

§ 3. *Experiments with Camphor.*

The formation of luminous pulses evidently requires, as Joubert recognised, that combination between oxygen and phosphorus vapour should not occur until a certain concentration of oxygen is reached. At this point in the investigation it appeared clear that this restraining action was due to a peculiar action of water vapour. One could not fail to be reminded of the well known action of ethylene, oil of turpentine, and numerous other substances, organic and inorganic, which are known to inhibit the oxidation and glow of phosphorus. Their action, it is true, is far more powerful than the action of water, but it seemed likely that the difference was only one of degree: if so, a judicious use of these substances would make it possible to get the travelling pulses in a much more striking form. This expectation was eventually confirmed. The earlier experiments made with ethylene and turpentine were not successful: I pass, therefore, to those made with the vapour of camphor:—

(1) A tubular bottle of 260 c.c. capacity, about 15 cm. long and 4·5 cm. diameter, was used. At the bottom were placed a few pieces of phosphorus and a few pieces of camphor, about the size of peas.

No precaution was taken to keep the camphor and phosphorus separate, nor did this seem necessary.[†] The bottle was exhausted, and air allowed to leak in slowly through a capillary tube. The general arrangement was similar to that of fig. 3, though in the present case the phosphorus was not suspended. The air pressure rose at a rate of about 0·28 mm. per minute.

Five minutes after the evacuation, a brilliant pulse or flash of greenish yellow light was seen to pass down the bottle. Another similar flash passed after 1½ minutes, and after that succeeded a series of flashes. The interval between them soon settled down to a pretty regular value of about 1 minute.

If the experiment is allowed to continue with the same rate of admission of air, nitrogen accumulates in the vessel and makes diffusion less easy. At a pressure of a few centimetres it is found that the luminosity which passes down lingers on the surface of the phosphorus for a second or two, and then for longer. Finally the phosphorus remains continuously aglow, and the pulses cease.

---

[*] 'Phil. Trans.,' 1888, p. 584.

[†] Though an apparent charring of the camphor was noticed after some weeks.

It is possible, however, to work in an atmosphere of nitrogen at atmospheric pressure. In this case the admission of oxygen must be slower to allow for the much longer time needed for the diffusion of phosphorus vapour away from the phosphorus surface.

The same bottle, with the phosphorus and camphor undisturbed, was filled to atmospheric pressure with nitrogen, and a capillary tube, 45 cm. long and 1 mm. diameter, attached to the neck, so as to allow of slow diffusion of atmospheric air into the bottle. This showed most striking effects. Thin and very bright pulses would start up and travel slowly along the length of the bottle. Sometimes they started in the middle, dividing into two, one of which travelled up and the other down. At other times a pulse would start from the top or from the bottom, and traverse the entire distance in one direction. These pulses were from 1 to 2 mm. thick (diameter 45 mm.) and travelled about 2 cm. per second. They contrast with the pulses observed in a rarified atmosphere (1 cm. pressure), which were 10 to 20 mm. thick, and travelled at a speed roughly estimated at 1 metre per second. The latter were much brighter, as is to be expected from the much shorter time they take to traverse the vessel, for the amount of phosphorus burnt in each case is not very different.

The thinness of the pulses at atmospheric pressure is due, no doubt, to the diminished facility of diffusion. Sometimes they were very flat, and accurately perpendicular to the axis of the cylinder in which they moved. At other times they had irregular curvature and were oblique to the direction of propagation. These differences are doubtless to be attributed to irregularities of temperature in the vessel.

In the apparatus as described, the average interval between pulses was about 8 minutes, though its action was liable to temporary disturbance by changes in the barometric height. It has remained in good order for several weeks at the time of writing, and there is no apparent reason why it should not remain so for much longer.

Some other substances give very similar results to camphor. They were tried in the exhausted bottle only. A piece of filter-paper, moistened with the liquid, was placed at the bottom, and a piece of phosphorus hung just above (as in fig. 3). The substances found to behave thus were nitrobenzene, butyl alcohol, amyl nitrite, and oil of bitter almonds—one drop only of the latter on filter paper. No doubt the list might be much extended. Ammonia may be added to it (see § 6).

§ 4. *Experiments with Pear Oil.*

A piece of filter paper was moistened with pear oil (consisting chiefly of amyl acetate), and placed at the bottom of a narrow bottle (height 18 cm., diameter 6 cm.), with a slice of phosphorus hung just above it. The bottle was exhausted, and air allowed slowly to leak in through a capillary tube. At first the phenomena were similar to those observed with camphor, but at a pressure of about 5 cm. a change occurred, the luminous pulse starting not from the top, where air was entering, but from the phosphorus surface. At the same time luminosity would cling to the phosphorus surface a second or two, after which it was extinguished. This tendency of the pulse to start from the phosphorus surface was repeatedly noticed in various experiments with pear oil under different conditions. Amyl alcohol gives the same effect. It is possible that camphor, or some of the other substances, which have been classified with it, might do the same if the right conditions were found. But they certainly do it less readily, if at all.

§ 5. *Oil of Turpentine.*

If a full quantity of this oil is placed in the bottle, luminosity is entirely stopped at all pressures up to the atmospheric. It was thought at first that a reduced quantity would probably give the brilliant luminous pulses observed with camphor, but this did not prove to be the case. When only one drop was placed on filter paper, and introduced into the bottle, there was still enough vapour to inhibit the oxidation of phosphorus at low pressures.

The one drop may still have been enough to give saturated vapour. In order to reduce the effect, turpentine was dissolved in 100 times its volume of olive oil, or (in another experiment) in the same volume of medicinal paraffin. It was then found possible, with a very slow feed of air, to get periodic luminous waves, but they were of a different character from the comparatively thin and very bright pulses obtained with, *e.g.*, camphor. The head of the luminous wave was equally definite, but it left the whole track luminous behind it for a perceptible interval, after it had reached the bottom. All then became dark again.

These experiments seem to show that turpentine is not radically different in its inhibiting action from substances like camphor. The difference that there is cannot, however, be explained as merely due to a stronger action;

for no concentration of turpentine, great or small, will give the effects as well as camphor.

The following oils, which were casually tried, seemed to behave like turpentine: aniseed, lavender, peppermint, eucalyptus, cinnamon.

§6. *Experiments with Ammonia.*

Ammonia gives effects like camphor. Its solubility in water, and the convenient control obtained by diluting the solution, makes it suitable for tests of how the period is controlled by the strength of the inhibiting substance, when air is admitted at a definite rate. The vessel used was a cylindrical bottle, 18 cm. long, and 6 cm. diameter. A capillary tube admitted air so as to raise the pressure 7 mm. of mercury per hour. The same capillary remained in used throughout the experiments with ammonia. A slice of phosphorus was hung up about 1 cm. above the layer of liquid at the bottom of the bottle (fig. 3).

Observations were made on the time required after evacuation for the pulses to begin, and also on the period. The latter was rather irregular at first, and somewhat longer than the steady value given, which was attained after the first few flashes had passed. Some residual irregularity remained, however, even when this comparatively steady state had been reached.

Ammonia density.	Time before first flash.	Period.
	minutes.	minutes.
[1·00*	1·1	0·083]
0·997	4·5	0·43
0·994	12·0	0·55
0·982	52·0	1·2
0·965	150·0	3·1
0·960	More than 240	4·1
0·955	Character of phenomena changed.	

* Water only.

Both the period and the time taken for the pulses to begin increase with the strength of the ammonia, but the latter time increases much more rapidly.

The time-lag before beginning measures the oxygen required to be present (1 minute = 0·023 mm. oxygen pressure), and the period measures on the same scale the oxygen consumed at each flash; for the steady *régime* implies that the oxygen consumed at each flash is the amount which leaks in during one period.

It appears, then, that the fraction of the whole stock-in-trade, so to speak, of oxygen, which is consumed at one flash diminishes when stronger ammonia is used, and the period becomes slower. With the weakest ammonia, about one-tenth of the whole amount present was consumed at each flash; with the strongest, less than one-sixtieth. It is not surprising, in the light of this, that the successive periods vary considerably in length. These periodic pulses, arising suddenly, travelling from one end of the vessel to the other, with completely dark intervals between them, are attributed to the combustion of the phosphorus vapour already present in the tube. Immediately after a pulse has passed, all the phosphorus vapour is consumed and a fresh quantity evaporates during the dark interval between the pulses.

Obviously, if this *régime* is to be followed, there is a limit to the period, under the condition of constant influx, as in the present experiments with ammonia. The period cannot be longer than the time required to admit enough oxygen to burn the whole of the saturated phosphorus vapour.

If we take the phosphorus molecule as P_4, oxidising to P_2O_5, it will combine with five times its volume, and hence with five times its partial pressure of oxygen. The careful measurements of Centnerszwer* give the vapour pressure of phosphorus vapour at 20° C. as 0·025 mm. This would require 0·125 mm. of oxygen for its complete combustion.

In the experiments with ammonia of longest period (4·1 minutes) the oxygen entering in a period is $4{\cdot}1 \times 0{\cdot}023$ mm. $=$ 0·094 m.m. This is ·75 of the amount required for complete combustion, and therefore it might be anticipated that the limiting period for simple pulses has practically been reached.

Experiment confirms this. When the ammonia concentration is increased still further it is found that the character of the *régime* changes entirely, though it still remains periodic. Thin pulses no longer travel down the vessel, and there are no longer completely dark intervals. The space round the phosphorus is filled with a diffuse luminosity, which increases periodically in brightness, but which is never reduced to zero and is not propagated through the vessel in a detached pulse right away from its place of origin. In this case the conditions which determine the period are altogether different. The quantity of phosphorus vapour burnt during a period is not limited by the saturation pressure, since the amount of evaporation while combustion is actually proceeding is no longer insensible.

---

* 'Zeit. Physik. Chemie,' vol. 85, p. 99 (1913).

§ 7. *Some Unclassified Observations.*

In addition to the phenomena described above, some others have been noticed which have not been investigated in detail. Thus in the experiments with camphor, I have sometimes seen a faint glow in the vessel before the passage of the brilliant pulse and heralding the approaching catastrophe. Sometimes too a faint glow remained for a second or two after the passage of the pulse.

Placing iodine in the vessel, the phosphorus surface would under some conditions light up and become abruptly extinguished again in a curious manner. It was thought that in this case chemical combination between phosphorus and iodine introduced some complication.

§ 8. *Probable Mode of Action of Inhibiting Substances in Preventing the Oxidation of Phosphorus.*

The power possessed by a trace of ethylene or turpentine or ammonia in hindering the glow of phosphorus has always seemed very strange and difficult to explain. It is sometimes said to be an example of "negative catalysis," but this phrase is not of much help.

Although no complete account can yet be given, the present experiments seem to show in what direction the explanation lies.

Consider a luminous wave travelling through a vessel containing phosphorus, air, and some inhibiting substance. When once this wave has been started, it is propagated through the vessel. This implies that combustion in any layer is started by contact with a previous layer which is already in the act of combustion. Now why does contact with the layer already in combustion produce this effect? In the case of a mixture of oxygen and hydrogen the reply would be that the high temperature developed by combustion heats the neighbouring layer and starts that into combustion also. In the present case such an explanation is scarcely tenable, for the rise of temperature to be expected is too small. Consider the propagation in the experiment of p. 377 when the vessel was filled with nitrogen at atmospheric pressure. Let us suppose that the phosphorus vapour was saturated and that the whole of it was burnt—the least favourable supposition for the argument. Taking the vapour pressure as 0·025 mm. (Centnerszwer, *loc. cit.*) and the molecule as P_4, a litre of nitrogen contains 0·181 mgrm. of phosphorus. The complete combustion of this quantity would give 1·07 calories, and this would only heat the containing atmosphere of nitrogen 3·6°. This then is the highest temperature which could be communicated to a neighbouring layer in which combustion had not yet begun. It may well be doubted whether so small a rise could determine the propagation of combustion: but there is evidence that propagation takes place under much less favourable conditions even than this.

We have supposed that in the experiments with camphor at atmospheric pressure, the full saturation quantity of phosphorus is being burnt. In the much less brilliant and more frequent pulses obtained in presence of water vapour, the quantity burnt must be much less—at a moderate estimate, *e.g.*, five times less. But we may go further still: for in some experiments (see p. 375) the pulses obtained in presence of water vapour, faint at the best, become much fainter still as they progress, and are reduced in intensity, perhaps five times further still, before they finally die out.

In this latter case, if we take the brightness as a measure of the rate of combustion, as it appears reasonable to do, we must suppose that perhaps not more than ·04 of the saturated concentration of phosphorus vapour is being consumed. This would only give a rise of temperature of about 0·14° C., and it is still harder to believe that this would determine propagation.

If propagation is not due to rise of temperature what is it due to? I think we must attribute it to the provision of nuclei.

There is a well known experiment on supersaturation which will illustrate what is meant. A supersaturated solution of, *e.g.*, sodium acetate, is contained in a long glass tube. If a fragment of the crystallised salt is dropped in at one end, crystallisation is propagated along the tube. The crystallisation of any given layer provides nuclei which determine the crystallisation of the next layer.

Similarly, I think we must suppose that the combustion of phosphorus provides nuclei which facilitate the action. The action postulated has some analogy with the action of platinum or other solid surfaces in promoting union of hydrogen and oxygen. The gaseous or particulate products of combustion of phosphorus are assumed to act somewhat similarly.

Coming now to the inhibiting substances, I think these must be supposed that these act by taking prior possession of the nuclei, and thus making them unavailable for propagation. That water, and still more, ammonia in presence of water, would seize on the products of oxidation of phosphorus, is no more than we may anticipate from the acid-forming character of the latter. In the case of camphor or turpentine, the kind of action here assumed has become familiar in studies of the mobility of gaseous ions in air at atmospheric pressure. The ions tend to unite with gaseous molecules, as evidenced by their diminished mobility in an electric field; and this effect is highly selective. A trace of water vapour or of an organic vapour has far more tendency to load the ions than a full atmosphere of nitrogen or hydrogen.*

The original spontaneous ignition is regarded as due to something which, from a molecular point of view, is a rare accident, concerning a few molecules only. Unless it can spread and multiply, it effects practically nothing: being in this respect like a single microbe in the putrefaction of wounds. The presence of a small quantity of a vapour like oil of turpentine, by attaching itself to the product, prevents the latter from inducing the combination of the neighbouring phosphorus molecules. From this point of view, it is not specially surprising that a very small quantity of the inhibiting substance should in some cases be enough to stop the action. It is often quoted as a surprising fact that one part of ethylene in 400 parts of air inhibits the glow of phosphorus. But, after all, even at this dilution, the partial pressure of the ethylene is still about 100 times the saturation pressure of the phosphorus vapour itself. There can be little doubt that the glow of phosphorus is in all cases due to the oxidation of vapour, since the phenomenon may be traced continuously from atmospheric pressure to low pressures, and in the latter case occurs visibly at a distance from the phosphorus surface. At low pressures the vapour has a chance to diffuse away to some distance before getting oxidised.

Where, as in the periodic luminous pulses, an increasingly favourable composition of the phosphorus-oxygen mixture eventually starts a wave of combustion in spite of the inhibitor, we must suppose that the action of the inhibitor in destroying the nuclei is overwhelmed by the too rapid multiplication of the latter; just as the too rapid multiplication of invading bacteria in a wound may overwhelm the destructive action of the phagocytes and enable infection to spread.

The question may be asked why some of the inhibiting substances—like water, camphor, and ammonia—lend themselves so much more readily to the formation of intermittent luminous pulses than others, like turpentine and eucalyptus oil. As already shown, the latter class can be made to give the effect by dilution, but not so well.

Striking as the difference is, it is probably not very fundamental—not perhaps more fundamental than the condition of "hunting" or not hunting in a self-regulating mechanism, such as a steam engine governor or an automatic arc lamp.

In these cases a general explanation of the action of the mechanism is by no means sufficient to determine whether it will "hunt" or not. Nor can the provisional theory of the action of inhibiting substances which has been given be expected to do so.

---

* To avoid misconception it is well to state specifically that I do not assume that the ions produced in the oxidation of phosphorus have necessarily anything to do with the matter.

§ 9. *Summary.*

The paper deals with the intermittent or periodic luminosity observed to be propagated through the gas space when the last traces of oxygen are being removed from air by means of phosphorus, or when air is allowed slowly to leak into an exhausted vessel containing phosphorus.

(1) It is shown that this effect, as ordinarily observed, requires the presence of water vapour. Moderate drying (*e.g.*, by sulphuric acid) makes the glow perfectly steady. Water vapour has therefore the power of inhibiting the combination of phosphorus vapour and oxygen within certain limits. When the composition of the mixture becomes favourable beyond those limits, a wave of combustion is propagated.

(2) Other substances are known to inhibit the glow of phosphorus. Some of these can be used to exhibit the above phenomena in a far more striking form than water. Camphor, ammonia, and pear oil are among the most effective, and experiments with them are described in detail.

(3) It is shown that the propagation of these waves of combustion cannot be attributed to the rise of temperature of one layer igniting the next layer, for the rise of temperature is too small. An alternative theory of the propagation is proposed, which assumes that it depends on the provision of nuclei, as in the propagation of crystallisation through a supercooled liquid.

(4) On this basis a theory of the action of the inhibitors (sometimes called "negative catalysts") is developed.

Some Experiments on Thermal Diffusion.

By T. L. IBBS, M.C., B.Sc., Assistant Lecturer in Physics, University of Birmingham.

With a Note on the Experiments by Prof. S. CHAPMAN, F.R.S.

(Communicated by Prof. S. W. J. Smith, F.R.S. Received March 16, 1921.)

Introduction.

It has been shown theoretically by Dr. S. Chapman* that a temperature gradient applied to a uniform mixture of two gases will tend to produce non-uniformity of composition, the heavier and larger molecules diffusing towards the cooler side, and the smaller and lighter molecules diffusing towards the hotter side. This phenomenon was termed "thermal diffusion." The difference in composition due to thermal diffusion increases until it is balanced by the opposite effects of ordinary diffusion, when a steady state will be reached. The effect will be greatest when the gases are mixed in nearly equal proportions by volume, and will be greater the more unequal are the masses and diameters of the gas molecules. It was also shown that the extent of the effect would vary with the character of the gas molecules, being at a maximum when the molecules behave like rigid elastic spheres. In the case where molecules behave like fifth power centres of force the effect would disappear entirely.

Experiments made by Chapman and Dootson† established the existence of the phenomenon. The results were chiefly qualitative, the nature and order of magnitude being in agreement with the theory. In no case were the differences equal to the theoretical values obtained on the assumption that gas molecules behave like rigid elastic spheres. Although the results are not claimed to be exact, the general effect could be regarded as rather less than a half of this calculated effect. The character of the gas molecule required to give this result would, however, be in agreement with that obtained for the actual gas molecules from the investigation of the variation of viscosity with temperature.

The method used by the writer in carrying out the following experiments depends on the use of the katharometer‡ for accurate gas analysis. The apparatus was devised shortly after the appearance of Dr. S. Chapman's paper in 1916, by Dr. G. A. Shakespear, who, when opportunity occurred in December, 1918, tested Chapman's predictions. The gases used were coal gas and also mixtures of carbon dioxide and hydrogen, and a wide range of temperature differences was tried. The results being satisfactory Dr. Shakespear suggested to the present writer that a more detailed investigation would be of interest. I am indebted to Dr. Shakespear for the use of the apparatus and for valuable advice and assistance in carrying out the work. The mixtures of gases used throughout these experiments were of hydrogen and carbon dioxide in definite proportions. These lend themselves particularly well to detection by the katharometer of any changes in composition due to thermal diffusion: the action of the instrument depending on the differences of thermal conductivity of various gases.

---

* 'Phil. Trans.,' A (1916).

† 'Philosophical Magazine,' March, 1917—"A Note on Thermal Diffusion."

‡ *Cf.* G. A. Shakespear, "Note on the Katharometer," 'Roy. Soc. Proc.,' vol. 97, p. 273 (1920).

Method of Experiment.

A katharometer (K_1) was first calibrated by using nine known mixtures of the two gases made up in small quantities for this purpose, the mixtures in turn being exposed to r_2, and the deflection on G_1 noted. The electrical arrangement, which is the normal one, is shown diagrammatically in fig. 1. A graph showing composition and galvanometer deflection was plotted (fig. 2), and by reference to this the composition of any other mixture of the gases can be determined by means of the katharometer.

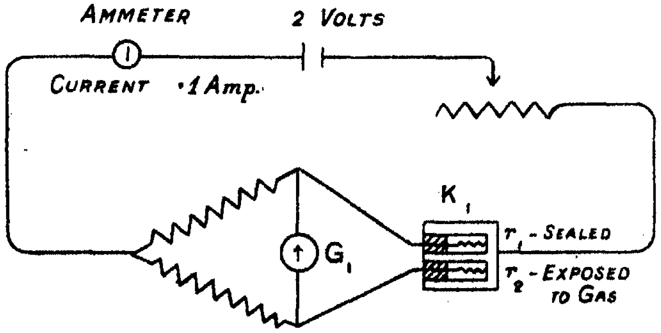


FIG. 1.—Electrical arrangement of katharometer K_1.

The mixtures for the actual experiments were required in considerable quantity, and they were made up under pressure in a 4 cubic feet cylinder. The composition of the contents of the cylinder was determined accurately in the way shown above.

A known mixture was passed from the cylinder through the apparatus shown in fig. 3. This apparatus, which can be termed the separator, consists essentially of a cylindrical glass vessel, down the middle of which runs a spiral of platinum wire. The spiral can be heated by passing a current through it. The gas mixture enters at A, passes upwards through the chamber B, and goes out either through C or D. Tube C takes the gas from

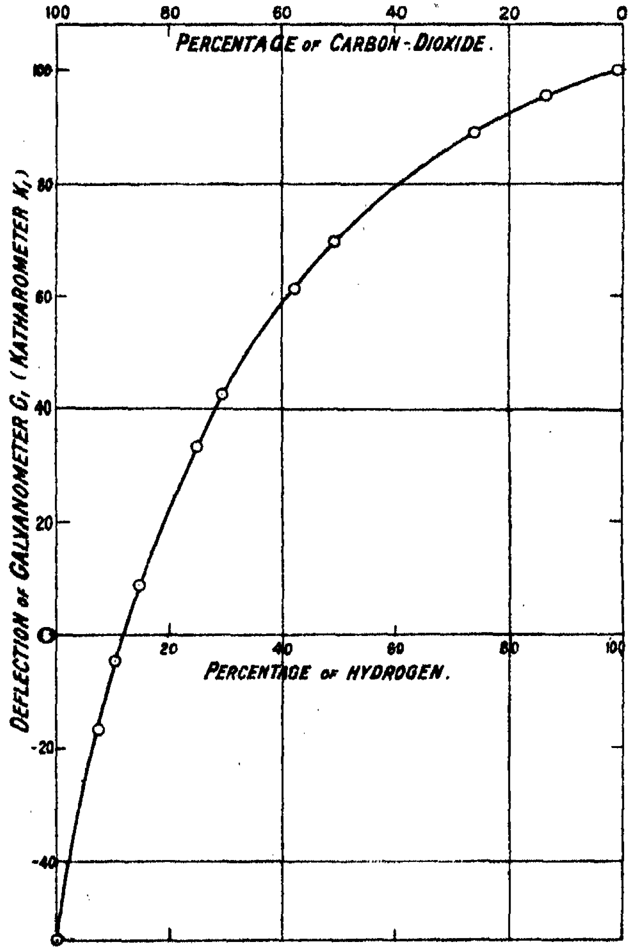


FIG. 2.

the cooler and outer part of the vessel, and tube D which surrounds the top of the hot spiral takes the gas from the middle and hotter part. During an experiment the gas was allowed to flow steadily through the apparatus. After passing out through C and D, the gas was passed through copper coils in order to equalise the temperatures of the two streams.

In order to detect and measure any difference in composition between the gases issuing from the two branches of the separator, use was made of a double flow or differential katharometer (K_2) through which the two streams of gas were passed.

This instrument is similar to the ordinary form except that in this case, instead of one cell being sealed and the other exposed to the gas, each cell communicates by small apertures with another cell exposed to the gas issuing from one branch of the separator. A difference in composition between the gases coming from the hot and cold sides will be detected by their difference in thermal conductivity. Fig. 4 is a diagrammatic representation of K_2, showing the two cells of the katharometer exposed to the gases. The electrical arrangement is the same as for K_1 (shown in fig. 1). Any difference in composition between the gases on the two sides of the katharometer upsets the balance of the bridge and causes a deflection on the galvanometer (G_2). In this case a reflecting galvanometer was used, greater sensitiveness being required. For a mixture of initial composition 50 per cent. hydrogen and 50 per cent. carbon dioxide, a percentage difference between the compositions of the gases on the two sides of less than 0·006 could be measured. For mixtures containing less hydrogen the instrument is more sensitive, and *vice versâ*. In the case of a mixture of 2 per cent. hydrogen and 98 per cent. carbon dioxide a difference of 0·001 per cent. could be measured, and with a

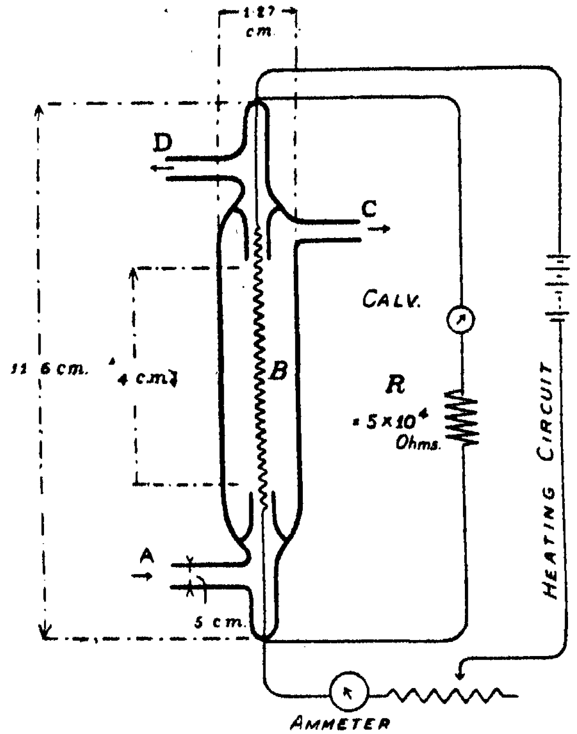


Fig. 3.—Separator with Electrical Connections.

mixture of 87 per cent. hydrogen and 13 per cent. carbon dioxide a difference of 0·01 per cent.

A calibration of K_2 was necessary to get the values of the differences in composition between the gases on the two sides, from the deflections of the galvanometer G_2. To effect this the apparatus shown in fig. 5 was used. For each mixture used in the experiments, a known small difference of composition was made in the following way. A gas mixture of known composition flows in at A, and divides into the two arms B and C. At E, in the arm C, a small quantity of hydrogen is introduced. This mixes in the chamber D, and the composition of the new mixture is determined by the katharometer K_1 (from the deflection of G_1—not shown in diagram). The mixtures in the two arms B and C now pass through the two sides of the differential katharometer K_2, the difference in composition causing a deflection of G_2 (not shown). Knowing the actual difference of composition in this case, we are able to get the meaning of a deflection of G_2 in terms of the difference in composition between the gases on the two sides of the katharometer.

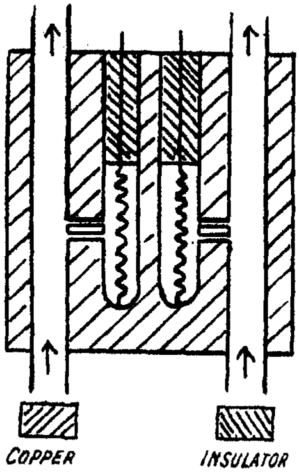


FIG. 4.—Differential katharometer, showing both cells exposed to gases.

The temperature of the platinum helix in the separator was varied by altering the current passing through it (fig. 3); the current being increased by regular amounts until the helix became red hot at about 600° C. The temperature gradient in the separator was from the temperature of the hot spiral down to that of the surrounding glass walls of the apparatus, *i.e.*, roughly, the room temperature (about 16° C.). No special cooling device

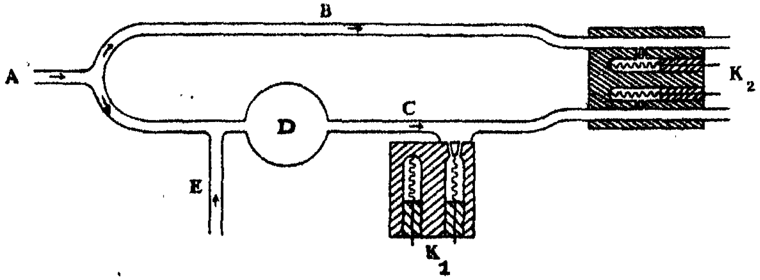


FIG. 5.

was used for keeping cool the outer walls of the apparatus, so that the real gradient of temperature in the separator was rather less than the calculated gradient. Thus it was found that, when the temperature of the heating spiral was increased, the amount of separation rose to a maximum, and then fell slightly, owing to the rise in temperature of the outer walls of the apparatus. For calculating the temperature of the spiral, it was necessary to know the potential difference between its ends, its resistance at a known temperature, and the temperature coefficient of resistance of the material of the wire, in addition to the current flowing through it. The potential difference was obtained by joining the ends of the spiral in series with a galvanometer and a resistance of 5×10^4 ohms (fig. 3), the galvanometer thus being used as a voltmeter. The calibration of this galvanometer and resistance as a voltmeter was made by a potentiometer method, using a Clark cell as a standard. The temperature coefficient of resistance for the platinum helix was found by measuring its resistance by a P.O. box at three known temperatures. A graph of temperature and resistance of the spiral was plotted by using a number of calculated values up to a temperature of 700° C., and, by reference to this, the temperature corresponding to any particular resistance could be obtained.

The method differed essentially from those of Chapman and Dootson, in that the thermal separation was continuous: a steady flow of the mixture being maintained in the separator, and the partially separated gases passing out on the hot and cold sides, the difference in composition being at once shown by the katharometer. This form of apparatus is well adapted for finding the effect of alteration of temperature gradient on the amount of thermal separation. The gases were allowed to pass through a water bubbler before finally escaping into the air. This served as a means of indicating the rate of flow through the apparatus, enabling a uniformly slow and steady flow to be maintained. The gas passed through each side of the bubbler at the same rate, which meant that approximately equal amounts of gas were being taken from the hot and cold regions of the separator. A steady state of separation (as indicated by galvanometer G_2) was reached after a mixture had been passing through the separator for about three minutes. The amount of separation obtained was not affected by the rate of flow through the separator, provided that the rate of flow was not made too great.

Results of Experiments.

Experiments were made on the following twelve different mixtures of carbon dioxide and hydrogen.

Table I.

Mixture No.	H_2.	CO_2.	Room temperature.
	per cent.	per cent.	° C.
1	2·0	98·0	12
2	7·7	92·3	13
3	10·7	89·3	13
4	19·7	80·3	17
5	24·2	75·8	18
6	38·0	62·0	18
7	41·5	58·5	17
8	50·0	50·0	16
9	55·5	44·5	16
10	61·0	39·0	16
11	72·5	27·5	16
12	86·5	13·5	17

In each case readings were taken with the platinum spiral of the separator varying between room temperature and red heat, thus giving temperature differences up to a maximum of about 600° C. The following Table of Results obtained from Mixture No. 10 may be regarded as a typical set:—

Table II.

Mixture No. 10. Composition: 61·0 per cent. H_2, 39·0 per cent. CO_2.
Room temperature 16·2° C.

Current in heating spiral.	Voltage between ends of spiral.	Resistance of spiral.	Temperature of spiral.	Percentage separation between gases from hot and cold sides.	Hence final composition.			
					Hot side.		Cold side.	
					H.	CO_2.	H_2.	CO_2.
amp.	volts.	ohms.	° C.	p.c.	p.c.	p.c.	p.c.	p.c.
0·05	0·82	16·40	25	0·045	61·022	38·978	60·978	39·022
0·10	1·66	16·60	35	0·157	61·078	38·922	60·922	39·078
0·15	2·52	16·80	45	0·344	61·172	38·828	60·828	39·172
0·20	3·52	17·60	80	0·600	61·300	38·700	60·700	39·300
0·25	4·60	18·40	115	0·891	61·445	38·555	60·555	39·445
0·30	5·71	19·00	145	1·22	61·61	38·39	60·39	39·61
0·35	7·00	20·00	190	1·56	61·78	38·22	60·22	39·78
0·40	8·42	21·05	240	1·92	61·96	38·04	60·04	39·96
0·45	10·10	22·30	300	2·29	62·14	37·86	59·86	40·14
0·50	11·76	23·52	365	2·61	62·30	37·70	59·70	40·30
0·55	13·51	24·55	420	2·84	62·42	37·58	59·58	40·42
0·60	15·56	25·98	500	3·31	62·65	37·35	59·35	40·65
0·65	17·62	27·20	570	3·52	62·76	37·24	59·24	40·76

The difference in composition due to thermal diffusion between the gases issuing from the hot and cold sides of the separator will, for the rest of this paper, be expressed simply as a percentage separation. In all cases the quantity of gas issuing from the two sides was practically the same, so that the actual composition of these gases can easily be found, if required, as in the last columns of the Table shown above. The effects of thermal diffusion were clearly observed in all the experiments, the excess of hydrogen always being found in the gas coming from the hot side and of carbon dioxide in the gas from the cold side. The results also show the increase of the effect with the increase of the temperature gradient.

To avoid giving large masses of data, the results are briefly summarised in

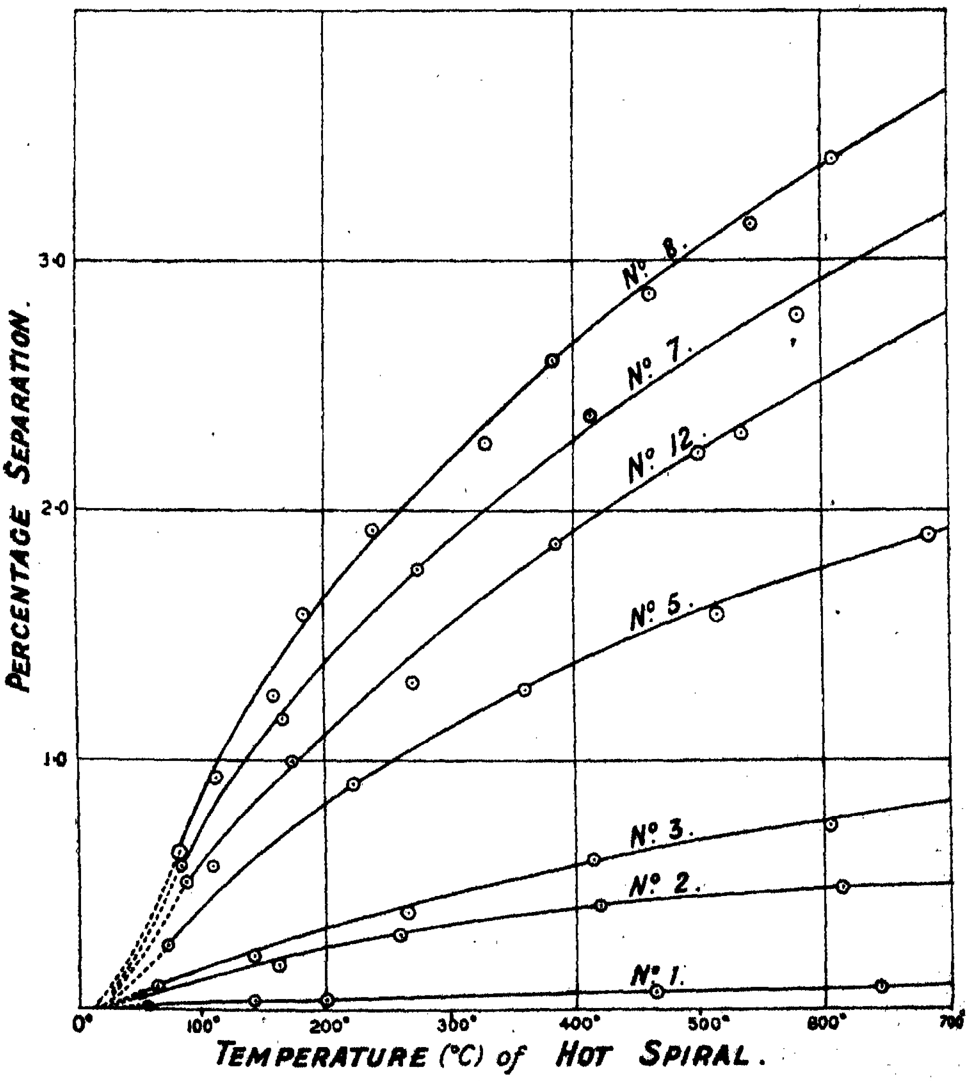


FIG. 6.—Temperature-separation curves for Mixtures 1, 2, 3, 5, 7, 8, 12. (See Table I.)

the form of curves showing the temperature of the hot spiral and the resulting percentage separation. Such curves for Mixtures Nos. 1, 2, 3, 5, 7, 8, and 12 are shown in fig. 6. Each curve represents a complete series of experiments.

By using such a series of graphs the relation between the amount of separation and the composition of the original gas mixture, for any definite temperature gradient, can be investigated. For example, we can pick out the different amounts of separation in each for a temperature difference of, say, 200° C.; and combining these with the corresponding composition of the original gas, construct another diagram. Such diagrams are shown in fig. 7 for temperature differences of 100°, 200°, 400°, and 600° C. It will be observed that these diagrams all have the same general form, the amount of separation increasing from zero to a maximum, and then diminishing to zero, in each case the maximum separation being shown for a mixture containing from 50 to 60 per cent. of hydrogen. This experimental result is in agreement with the theory of Chapman, who showed that the maximum separation should occur in mixtures containing about 50 per cent. of each gas. Although the general form of the curves is apparent the points do not lie as uniformly as those on the temperature-separation curve for a given mixture. This may be expected from the fact that the points are obtained from different sets of experiments, in which the only check on getting an equal flow of gas from the hot and cold sides was by passing it through a water-bubbler.

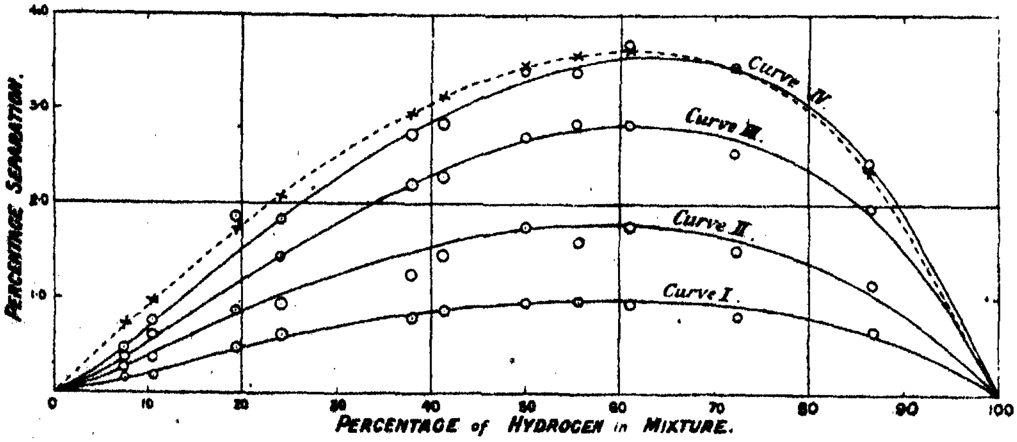


FIG. 7.—Relation between composition and separation with a given temperature difference.

Curve I—100° C. difference; Curve II—200° C. difference; Curve III—400° C. difference; Curve IV—600° C. difference; Dotted Curve—Obtained from values of Ck_t calculated.

Analysis and Summary of Results.

The following relation was deduced theoretically by Chapman. If the absolute temperature is denoted by T, and the proportion by volume of the heavier and lighter gases by v_1 and v_2 respectively (so that $v_1+v_2 = 1$), then in the steady state the concentration and temperature gradients at a point $(x \,.\, y \,.\, z)$ are related to one another by three $(x \,.\, y \,.\, z)$ equations of the form

$$\frac{d}{dx}\left(\frac{v_1}{v_1+v_2}\right) = -\frac{d}{dx}\left(\frac{v_2}{v_1+v_2}\right) = k_t\,\frac{d}{dx}(\log \mathrm{T}),$$

where k_t is the ratio of the ordinary coefficient of diffusion to the coefficient of thermal diffusion. Consequently it follows that over a distance within which k_t may be regarded as constant, v_1/v_1+v_2 varies as log T.

A further series of graphs have been plotted showing percentage separation and $\log_{10} \mathrm{T}_1/\mathrm{T}_2$, where T_1 is the absolute temperature of the spiral and T_2 is the room temperature. These graphs for Mixtures Nos. 2, 4, 6, 10, and 12 are shown in fig. 8. The general form of these graphs approximates closely to a straight line when the temperature differences exceed about 100° C. The slight irregularity near the origin may be due to the form of the heating spiral. The nature of these graphs may be regarded as constituting a striking experimental verification of the theory.

The greatest percentage separation obtained in the experiments was 3·52 for mixture No. 10 (61 per cent. H_2 and 39 per cent. CO_2), with a temperature difference of 557° C. Chapman and Dootson obtained a greatest percentage separation of 3·6, using the same gases (43·1 per cent. H and 56·9 per cent. CO_2), with a temperature difference of 220° C. The smaller amount of separation is to be expected from the form of apparatus used, in which the area of the hot spiral is small compared with that of the outer glass walls of the separator. The simpler form of separating device employed by Chapman and Dootson facilitates calculations involving temperature gradient and the absolute amount of separation obtained. The shape of the heating spiral, and the fact that there is a continuous current of gas through the apparatus, make it impossible to calculate the configuration of the gas at the higher temperature. In this sense the results are comparative only: we cannot obtain a value for K_t and compare it with values obtained theoretically. The results of the experiments, therefore, do not furnish any information as to the nature of the molecular model which would most nearly represent the actual gas molecules. They do, however—

(*a*) Provide a verification by a new method of part of the results obtained by Chapman and Dootson.

(*b*) Show that the maximum separation due to thermal diffusion is obtainable in a mixture containing 50–60 per cent. of the lighter gas.

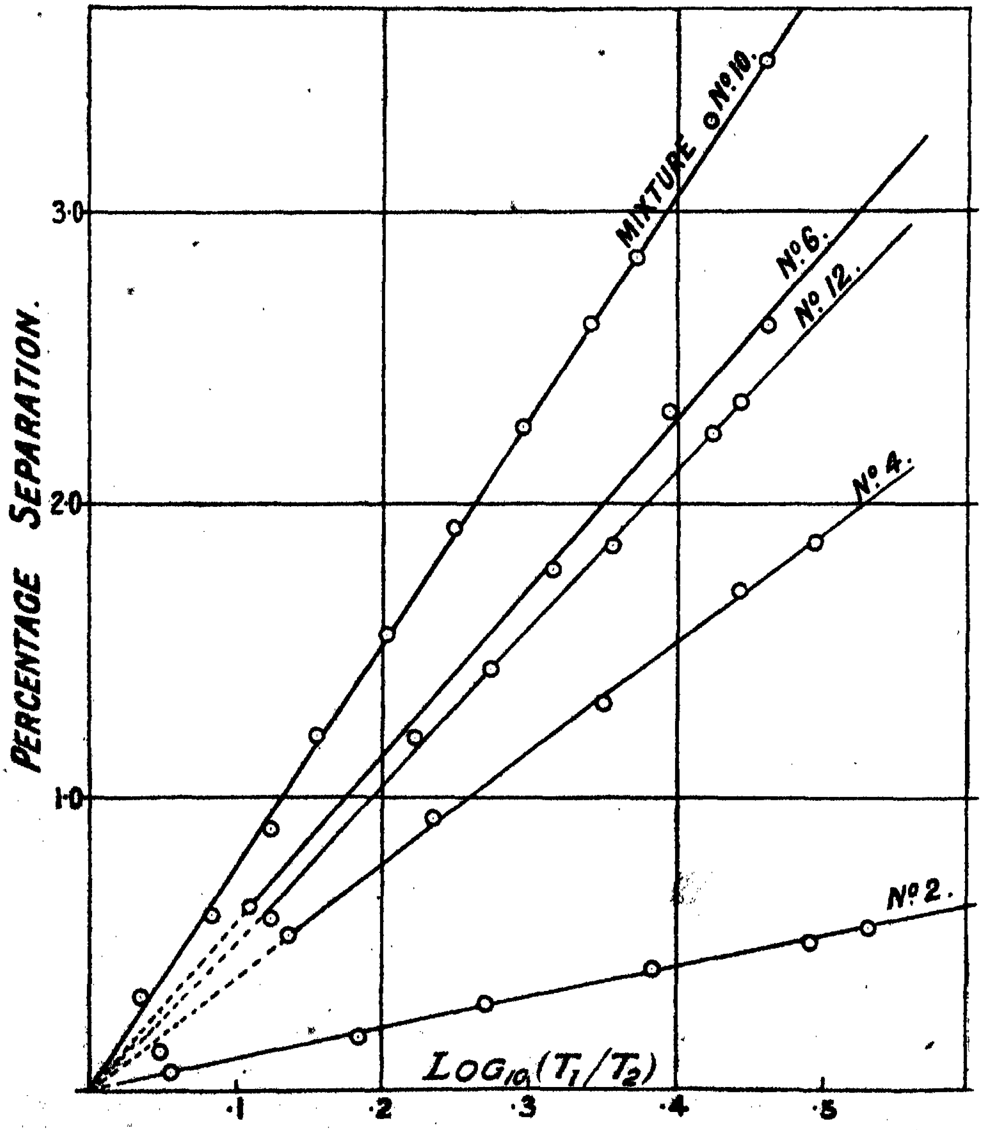


FIG. 8.—Relation between percentage separation and $\log_{10} T_1/T_2$ for Mixtures 2, 4, 6, 10 and 12.

(*c*) Show that the amount of separation varies as $\log T_1/T_2$, where T_1 and T_2 are the absolute temperatures forming the temperature gradient.

Note on Mr. T. L. Ibbs' Experiments.

By Prof. S. Chapman, F.R.S.

Mr. Ibbs' interesting and accurate experiments may advantageously be compared, in rather more detail than he has done, with the theoretically predicted influence of the ratio of the two gases in a mixture upon the coefficient of thermal diffusion k_t. It is sufficient to consider the simple formula giving the first approximation to k_t (correct to within a few per cent.) for rigid elastic spherical molecules. The value of k_t for actual gas molecules is probably only about half that calculated for this molecular model, but the variation with the ratio of mixture is likely to be nearly the same in the two cases.

The formula in question, with numerical values inserted for the H_2–CO_2 mixture, is*

$$k_t = \frac{0{\cdot}331\lambda_1 + 0{\cdot}0542\lambda_2}{47{\cdot}2 + 86{\cdot}5\lambda_{12} + 5{\cdot}58\lambda_{21}};$$

λ_1, λ_2 are the percentages, by volume, of carbon dioxide and hydrogen respectively, while $\lambda_{12} \equiv \lambda_1/\lambda_2$, $\lambda_{21} \equiv \lambda_2/\lambda_1$. The following values are calculated for the twelve mixtures used by Mr. Ibbs:—

Mixture.	1.	2.	3.	4.	5.	6.
k_t	0·008	0·029	0·039	0·069	0·082	0·118
Ck_t	0·2	0·72	0·97	1·72	2·05	2·95

Mixture.	7.	8.	9.	10.	11.	12.
k_t	0·125	0·138	0·144	0·146	0·138	0·095
Ck_t	3·12	3·45	3·60	3·65	3·45	2·37

The last row is added to facilitate comparison with the uppermost curve in fig. 7, with which the curve for k_t is, as it should be, very closely similar.

As regards the absolute amount of separation found by Mr. Ibbs, it is sufficient to consider mixture No. 10, for which fig. 8 shows a separation of 3·85 per cent. when $\log_{10} T_1/T_2 = 0{\cdot}5$. The theoretical amount for rigid spherical molecules is $0{\cdot}146 \cdot \log_e 10 \cdot 0{\cdot}5$, or 16·8 per cent.; the observed value is much less, partly because of the molecular model above used, which makes k_t about twice too* great; and, secondly, because T_1/T_2 measures the ratio of the extreme temperatures in the separator, and not the ratio of the mean temperatures of the two halves of the gas.

---

* 'Phil. Mag.,' vol. 33, p. 522 (1917).

A point of special interest in Mr. Ibbs' paper is the shortness of the time, about three minutes, required for thermal diffusion to take effect.

The Dynamics of Revolving Fluid on a Rotating Earth.

By Capt. D. Brunt, M.A., B.Sc.

(Communicated by Sir Napier Shaw, F.R.S. Received April 6, 1921.)

In a paper on the dynamics of revolving fluids, the late Lord Rayleigh* considered the special case of fluid revolving about a fixed vertical axis, neglecting the rotation of the earth. The object of the present paper is to investigate the modifications of Rayleigh's results which are brought about by the rotation of the earth, and by translation in a vertical plane of the axis of symmetry. Air is treated as an incompressible non-viscous fluid.

Let the motion be referred to rectangular axes, x, y, z, rotating with the earth, the axis of z being vertical and the axes of x and y in the horizontal plane. At a point x, y, z, the components of velocity are u, v, w, the pressure is p, and density ρ. If gravity is the only impressed force, the equations of motion are

$$Du/Dt = -1/\rho \; dp/dx + lv, \tag{1}$$

$$Dv/Dt = -1/\rho \; dp/dy - lu, \tag{2}$$

$$Dw/Dt = -1/\rho \; dp/dz + g, \tag{3}$$

where $$D/Dt = d/dt + u\, d/dx + v\, d/dy + w\, d/dz$$

and the equation of continuity is

$$du/dx + dv/dy + dw/dz = 0, \tag{4}$$

where $l = 2\omega \sin \phi$, ω being the angular velocity of rotation of the earth, and ϕ the latitude. The terms lv, $-lu$, in the first two equations represent components of the deviating force due to the earth's rotation, and the inclusion of these terms takes complete account of the rotation of the earth.

If values of u, v, w, p, be found to satisfy equations (1) to (4), then if U be a constant, $u+U$, v, w, $p-\rho l U y$ will also satisfy these equations. In other words, to any system of velocities which satisfies the equations there may be added a uniform velocity of translation U, provided the pressure distribution be corrected by the addition of a horizontal gradient of pressure of amount $2\omega\rho U \sin\phi$ at right angles to U. This affords a simple method of proceeding from the case of fluid rotating about a fixed axis to the case where the axis has a uniform velocity of translation, and we may therefore restrict the discussion, in the first place, to the case of motion symmetrical about a fixed axis.

---

* 'Roy. Soc. Proc.,' A, vol. 93, p. 148 (1916).

It is convenient to refer the motion to cylindrical co-ordinates, r, θ, z, the velocities u, v, w being reckoned positive in the directions of r, θ, z, increasing. It is assumed that the motion is symmetrical about the axis of z. The only impressed force being gravity, the equations of motion are

$$du/dt + u\,du/dr - v^2/r + w\,du/dz = -1/\rho\,dp/dr + lv, \tag{5}$$

$$dv/dt + u\,dv/dr + uv/r + w\,dv/dz = -lu, \tag{6}$$

$$dw/dt + u\,dw/dr + w\,dw/dz = -1/\rho\,dp/dz + g. \tag{7}$$

In the special case where $u = w = 0$, equation (5) reduces to the ordinary equation for the gradient wind for circular isobars.

The case where $w = 0$, and u and v are functions of r and t only, is the most interesting case considered by Rayleigh. The motion is then two-dimensional, and he expresses it conveniently by the vorticity ζ which moves with the fluid, and the stream-line function ψ, which are connected by the equation

$$1/r\,d/dr\,(r\,d\psi/dr) + 1/r^2\,d^2\psi/d\theta^2 = 2\zeta, \tag{8}$$

and

$$u = -d\psi/r d\theta, \tag{9}$$

$$v = d\psi/dr. \tag{10}$$

Since u is independent of θ, $d^2\psi/d\theta^2 = 0$, and

$$1/r\,d/dr\,(r\,d\psi/dr) = 2\zeta.$$

The solution of these two equations therefore yields

$$\psi = 2\int dr/r \int \zeta r\,dr + A\log r + B\theta, \tag{11}$$

where A and B are independent of r and θ, but may vary with time t. We then have

$$u = -B/r, \tag{12}$$

$$v = 2/r\int \zeta r\,dr + A/r. \tag{13}$$

In the case where the fluid has an initial rotation as a solid body (with angular velocity ζ) there is a ζ initially, and therefore permanently, constant throughout the fluid, and equation (13) reduces to

$$v = \zeta r + A/r. \tag{14}$$

The effect of convergence towards the axis is therefore to superpose upon the rotation as a solid body a simple vortex of intensity, A.

Equations (8) to (14) are given by Rayleigh, and are independent of any consideration of the rotation of the earth. The rotation of the earth only affects the determination of the value of A at any given time.

Equation (6) may be written

$$(d/dt+u\,d/dr+w\,d/dz)\,(rv) = -lru = l\mathrm{B}. \tag{15}$$

This equation expresses the rate of change of the product rv for a particular ring of particles.

The rotating fluid may be supposed to be contained within a cylindrical wall, which moves inward as the fluid converges, fluid being removed along the axis, without disturbing the horizontal motion. Let R_0 be the radius of the wall at time $t = 0$ and R the radius at time t. It has been supposed that initially the fluid rotates as a solid body, so that initially $\mathrm{A} = 0$.

Equation $u = -\mathrm{B}/r$ may be written

$$r\,dr/dt = -\mathrm{B},$$

and on integration this yields

$$\int_0^t \mathrm{B}\,dt = -\tfrac{1}{2}\Big[r^2\Big]_0^t = \tfrac{1}{2}(\mathrm{R}_0^2-\mathrm{R}^2). \tag{16}$$

Integrating equation (15) and substituting from (16) we find

$$(rv)_t-(rv)_0 = l\int_0^t \mathrm{B}dt = \tfrac{1}{2}\,l\,(\mathrm{R}_0^2-\mathrm{R}^2). \tag{17}$$

This equation could also have been derived from the constancy of the circulation round a circuit made up of the same particles. Taking account of the fact that the xy plane has a rotation $\omega\sin\phi$ about the axis of z, the circulation round a circuit of radius r is $2\pi\,(rv+r^2\,\omega\sin\phi)$, whence equation (17) immediately follows.

But from equation (14), we have, for the boundary,

$$(rv)_0 = \zeta\mathrm{R}_0^2, \qquad (rv)_t = \zeta\mathrm{R}^2+\mathrm{A},$$

$$\begin{aligned}\therefore\qquad \mathrm{A} &= (rv)_t-(rv)_0+\zeta(\mathrm{R}_0^2-\mathrm{R}^2)\\ &= (\zeta+\tfrac{1}{2}l)(\mathrm{R}_0^2-\mathrm{R}^2),\\ \mathrm{A} &= (\zeta+\omega\sin\phi)\,(\mathrm{R}_0^2-\mathrm{R}^2).\end{aligned} \tag{18}$$

It follows from (14) that

$$v = \zeta r+\frac{(\zeta+\omega\sin\phi)\,(\mathrm{R}_0^2-\mathrm{R}^2)}{r}. \tag{19}$$

Equation (19) differs from Rayleigh's* equation (19) only by the inclusion of the term $\omega\sin\phi$, which represents the effect of the rotation of the earth.

---

* *Loc. cit.*, p. 153.

The effect of the earth's rotation is to increase the strength of the superposed simple vortex. It is clear that, when there is originally no rotation, or $\zeta = 0$, the effect of convergence due to the removal of air is to produce a simple vortex, in which the distribution of velocity is given by

$$v = \frac{\omega \sin \phi \, (R_0^2 - R^2)}{r}. \tag{20}$$

This equation thus gives the effect of the removal of air in a region where there is no vorticity, in contrast with equation (19), which shows the effect in a region having originally uniform vorticity, ζ.

If a cyclone starts as a uniformly rotating disc of radius R_0, with uniform angular velocity ζ, the effect of removal of fluid along the axis is to cause the outer radius R_0 to shrink to a smaller value, R, and, within the region bounded by the cylinder of radius R, to superpose upon the uniform rotation a simple vortex whose intensity increases steadily as the removal proceeds (equation (19)). Outside the cylinder of radius R, the change of motion caused by convergence is represented by a simple vortex, in accordance with equation (20).

The amount of fluid which must be removed from a layer of the vortex of unit thickness, in unit time, is $-2\pi r u$ or $2\pi B$. In a time t the amount of fluid removed is $2\pi \int_0^t B dt$ or $\pi (R_0^2 - R^2)$, by equation (17), showing that the intensity of the simple vortex produced by the convergence is directly proportional to the total amount of fluid removed from the region of the axis, and *vice versâ*, the removal of that amount of fluid implies the development of the vortex.

The distribution of pressure is evaluated by the integration of equation (5). If at any time the u motion ceases, the subsequent distribution of pressure is obtained by integrating

$$1/\rho \; dp/dr = v^2/r + lv = v^2/r + 2\omega \sin \phi \, v,$$

or, from (19)

$$1/\rho \; dp/dr = \zeta(\zeta + 2\omega \sin \phi) r + \frac{2(\zeta + \omega \sin \phi)^2 (R_0^2 - R^2)}{r} + \frac{(\zeta + \omega \sin \phi)^2 (R_0^2 - R^2)^2}{r^3},$$

whence, neglecting the variations of ϕ,

$$p/\rho = \tfrac{1}{2} \zeta (\zeta + 2\omega \sin \phi) r^2 + 2(\zeta + \omega \sin \phi)^2 (R_0^2 - R^2) \log r - \tfrac{1}{2} \frac{(\zeta + \omega \sin \phi)^2 (R_0^2 - R^2)^2}{r^2} + \text{constant}. \tag{21}$$

The effect of the inward motion upon the distribution of pressure, while the inward motion is still in progress, is easily evaluated by substitution of $u = -\mathrm{B}/r$ in equation (5). Neglecting $w\,du/dz$, we find

$$1/\rho\; dp/dr = v^2/r + lv + \mathrm{B}^2/r^3 + 1/r\; d\mathrm{B}/dt,$$

which, on integration, yields the same result as equation (21), with the addition of $-\tfrac{1}{2}\,\mathrm{B}^2/r^2 + d\mathrm{B}/dt \log r$ to the right-hand side. The second of the additional terms vanishes if the rate of removal of fluid is constant.

The results derived above may be applied to the supply of air at the axis, involving a progressive increase of the extreme radius R of the disc of fluid concerned. The dynamical effect is to superpose on the original distribution of velocity a simple vortex of negative intensity, or a vortex rotating clockwise. A region of descending air should therefore, merely on account of its descent, acquire a clockwise rotation, with velocity inversely proportional to the distance from the centre of the axis of revolution. If one deals with a stratum of unit thickness, the added air might come from above or below. It is only a question of there being an addition of air to the layer.

If the fluid is added at the centre of the disc, so that the outer radius R_0 increases to R at time t, while the added fluid is bounded by a circle of radius R' at time t, equation (19) still holds within the limits $\mathrm{R}' < r < \mathrm{R}$, and equation (20) still holds for $r > \mathrm{R}$. But R is now greater than R_0, so that the simple whirl produced by divergence from the axis rotates clockwise. Within the added fluid, where $r < \mathrm{R}'$, it follows from the constancy of the circulation round a circle of particles (initially zero when the fluid is added at $r = 0$), that

$$rv + r^2\omega \sin\phi = 0,$$

or

$$v = -r\omega \sin\phi. \tag{22}$$

In other words, the added fluid forms a core rotating clockwise as a solid, with angular velocity $\omega \sin\phi$.

The stream-lines of instantaneous flow in the general case, where the distribution of velocity is represented by

$$u = -\mathrm{B}/r, \qquad v = \zeta r + \mathrm{A}/r$$

are obtained by integrating the equation

$$r\,d\theta/dr = v/u = -(\zeta r^2 + \mathrm{A})/\mathrm{B}. \tag{23}$$

This yields on integration

$$re^{\zeta r^2/2\mathrm{A}} = \mathrm{C}e^{-\mathrm{B}/\mathrm{A}\,\theta}, \tag{24}$$

and the stream-lines are spirals whose angle of convergence increases towards the centre, except in the special case where $\zeta = 0$, and the motion is equivalent to a simple vortex, in which case the stream-lines are equiangular spirals.

In the complex case, where the distribution of velocity is a simple vortex superposed upon solid rotation, the angle of convergence of the horizontal projection of the spiral paths cannot be constant over the whole area unless air is removed over the whole area. The relation $vr =$ constant is the only one which is consistent with air passing horizontally across the area, with no extraction vertically except at the centre.

In all that precedes, the centre of the two-dimensional whirl has been assumed at rest. If the whirl is given a uniform velocity of translation, U, the distribution of velocity relative to the centre remains unaltered, and the spiral paths of instantaneous flow relative to the centre of the whirl are unaltered. All that is necessary is to add a correcting term, $-2\omega \sin\phi \,.\, Uy$ to the equation (21) for the distribution of pressure.

Some Experiments on the Catalytic Reduction of Ethylene to Ethane.

By Dorothy Muriel Palmer, Girton College, Cambridge, and William George Palmer, St. John's College, Cambridge.

(Communicated by Sir Wm. Pope, F.R.S. Received April 14, 1921.)

Finely divided reduced nickel has been applied as an active agent in the catalysis of a great range of reactions, but very little is known about the mechanism of its action. It was thought that the intensive study of some particular process might be of interest, in view of the paucity of quantitative information about catalysis.

The reaction selected for investigation was the combination of ethylene and hydrogen to form ethane. Sabatier* states that in the presence of freshly reduced nickel this reaction begins at 30°–45° C., and continues with the evolution of heat; thus there are few of the difficulties encountered when working at higher temperatures. Ethylene and hydrogen can both be obtained in a high degree of purity, and since the reacting and resultant substances are gases with simple volume relations between them, it was found possible to watch the progress of the reaction by a static method.

This was done by observing the pressure changes produced in a closed vessel containing the catalyst, together with a mixture of ethylene and hydrogen of known composition.

---

* 'Comptes Rendus,' vol. 124, p. 1360.

The reaction chamber was a glass tube, heated electrically. It contained a glass ball which could be caused to roll from end to end of the tube by tilting the latter automatically at intervals, by this means an intimate mixture of the reacting gases was secured during the whole progress of the reaction. Unfortunately, this arrangement gradually forced the whole of the catalytic powder to the ends of the tube. In the series of experiments here described the catalyst was deposited on pellets of China clay, which were themselves, in the place of the glass ball, used to stir the gases. One end of the tube was sealed, the other was closed by a glass tap leading through a T-piece to a manometer, a gas reservoir, and a Fleuss pump, all isolated by suitable taps.

In the gas reservoir was a mixture of equal volumes of ethylene and hydrogen. The action takes place in accordance with the equation

$$C_2H_4 + H_2 = C_2H_6,$$

so that two volumes of such a gas mixture give one volume of ethane. The relative proportions of the gases in the reaction tube were kept constant by introducing this equal-volume mixture from the reservoir as fast as contraction took place. The total pressure inside the tube was thus maintained at a constant value; and a comparison of the rates of reaction was obtained by noting the rates at which it was necessary to pass gas mixture from the reservoir into the tube in order to keep the pressure constant.

The volume passed per minute was taken as the rate of the reaction. The rates so obtained were plotted against the times in minutes from the first observation, which was as soon as possible after the introduction of the gas mixture into the reaction tube, but not exactly the moment of introduction.

Two series of results were obtained. In one the temperature was nearly constant, the percentage of ethylene varying; in the other the temperature varied, but the reacting-mixture always contained equal volumes of the two gases. The limits of temperature within which observations were possible were 75°-120° C. The curve produced was always of the same form, showing a preliminary period in which the rate was negligible, then a very rapid increase in the rate to a maximum value, followed by a more gradual decrease. The position and height of the peak portion of the curve depended partly on the temperature and partly on the composition of the gas mixture.

The variations in temperature and composition were, of set purpose, made in haphazard order, and it was found that any curve could be repeated, even after a lapse of several weeks.

The results obtained with nickel powder and with the impregnated China clay pellets were similar, but the determinations with the powder were made with apparatus of a cruder form than that finally adopted; the curves are, therefore, not shown on the appended diagram. It is certain, however, that the China clay of the pellets played no part as catalyst.

Control experiments were performed with pure hydrogen and pure ethylene in the reaction tube. With ethylene there was no absorption whatever, but with hydrogen there was a small constant absorption, unchanged in rate after 2½ hours.

Preparation of the Catalyst and Reacting Gases.

Pure nickel oxide was prepared as follows. Sheet nickel was cleaned, then dissolved in pure dilute nitric acid. To this solution a slight excess of ammonia was added, the mixture was then boiled and filtered, to remove a small deposit of iron hydroxide. After a repetition of this process, there was no further formation of precipitate. The solution was then condensed, and excess of concentrated ammonia solution was added. Light violet crystals of a nickel-ammonia complex formed, were filtered off, and washed with concentrated ammonia. On boiling with much water, these crystals decomposed, precipitating a bright green hydroxide of nickel in a powdery form. This was collected, washed and dried, and, when needed, was dissolved in dilute pure nitric acid.

Cylindrical rods of china clay, of diameter 3 mm., were used as support for the catalyst. The composition of these rods was as follows:—

Per cent.
$SiO_2 = 65·27$
$Al_2O_3 = 30·06$
$Fe_2O_3 = 0·65$
$MgO = 0·173$

Per cent.
$CaO = 0·43$
$K_2O = 1·61$
$Na_2O = 0·723$

They were rendered inactive as catalysts by ignition to bright redness for four hours. Two or three of these rods were broken into short lengths, then soaked and boiled in the liquid obtained by concentrating the solution of the purified nickel hydroxide. The pellets were dried at 100° C., then gently ignited, shaken free of loose powder, and placed in the reaction tube. The deposit of nickel oxide was reduced by the passage through the tube of dry electrolytic hydrogen. At the temperature employed, 300° C., this reduction was very slow, and it was several days before water could no longer be detected in the issuing gases. The pellets were then allowed to cool in an atmosphere of hydrogen. No air was ever admitted into the reaction tube. The exit was sealed off when the reduction of the nickel was complete.

The total quantity of nickel in the reaction tube was estimated electrolytically at the end of the series of readings. Its quantity was found to be 0·068 grm.

Ethylene was prepared by the action of zinc-copper couple on freshly distilled ethylene dibromide, boiling at 131° C. The gas was washed by being bubbled through tubes containing more of the couple immersed in alcohol. A gas analysis by Hempel's method, using bromine water, indicated an ethylene-content of over 99 per cent.; but, in order to remove any possible traces of air or hydrogen, the ethylene was solidified by cooling in liquid air, and then fractionated. It was stored in an aspirator over water saturated with pure ethylene.

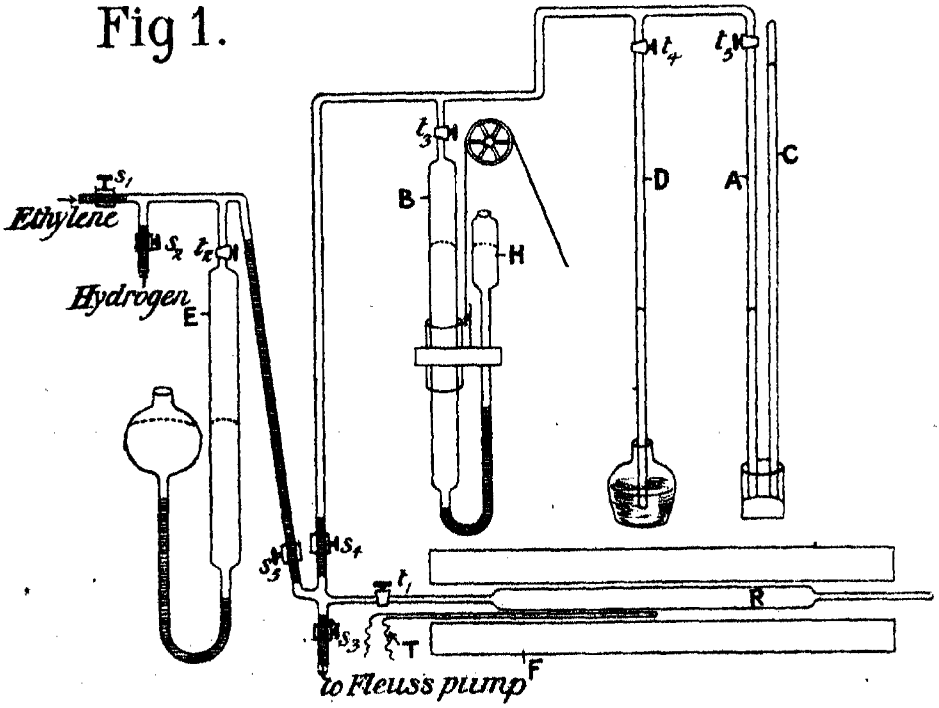
Hydrogen was prepared by the electrolysis of caustic potash, as described by one of us.[*] It also was stored over water.

When required, these gases were drawn through trains of soda-lime and calcium chloride tubes into the mixing vessel.

Apparatus.

The furnace, F, was especially constructed, so that the heating might be as even as possible throughout the cavity. The inner tube of brass was wrapped in asbestos paper, then evenly wound with a heating coil of constantan wire.

Fig 1.



This was insulated with asbestos cord, then the whole was packed in a cylindrical iron case filled with ignited magnesia. The furnace was supported at its middle point, so that it was free to turn in a vertical plane, and could be rocked up and down, either by hand or by an electric motor. The temperature inside the brass tube was read by means of an insulated thermocouple, T, with its hot junction placed at the middle of the tube, and its cold junction in a Dewar flask filled with melting ice (see fig. 1).

---

[*] 'J. Soc. Chem. Ind.,' vol. 39 (20), 316T.

The capacity of the reaction tube, R, was about 70 c.c. Its length was 30 cm., and its inside diameter 1·7 cm. The length was several inches less than that of the furnace, that there might be no cooling effect at the ends. When the prepared pellets had been placed inside, the ends of the tube were drawn out and fused to pieces of capillary tubing. The exit was sealed off after the reduction of the nickel was complete, and the entrance was closed by a glass tap, t_1. The volume of gas in these capillary end pieces was negligible. Asbestos wool was used to pack the glass reaction tube securely into the heating tube; this also prevented the entrance of air currents.

A plan of the apparatus is shown in the diagram. The shaded portions of the connections were made of rubber pressure tubing, the rest was glass; glass could not be used throughout, as flexible connections were necessary, to allow for the motion of the furnace. The taps t_1, etc., were of glass, those marked s were ordinary screw clips.

E, a gas burette filled with mercury, was used for the preparation of the required mixture of dried gases. A was a manometer dipping into mercury with an ordinary barometer C at its side. D was a water manometer, to read small alterations of pressure. It was isolated by the tap t_4 with capillary tubing on either side of it. As t_4 was opened only when the pressures on either side were approximately equal, there could be no appreciable escape of water vapour into the rest of the apparatus. Millimetre scales were fixed behind A, C and D.

B, a glass measuring tube graduated in tenths of a cubic centimetre, was filled with a mixture of equal volumes of ethylene and hydrogen. A small pulley and windlass allowed a very gradual raising of the mercury reservoir H; thus, the taps t_1, t_3, t_4 and s_4 being open, the liquid in D could be kept at a constant level by an adjustment of the height of H.

Procedure.

The furnace was heated overnight, to secure equilibrium of temperature inside the reaction tube. That constancy of temperature could be obtained was shown by a blank experiment with nitrogen in R. When the taps t_1, t_3, t_4, t_5 and s_4 were open, the liquid in D showed only small oscillations during a four-hour period of observation. B was filled with its equal volume mixture, and the gas mixture to be investigated was prepared in E.

The taps s_1, s_2, t_2, t_3, t_4, were closed, the others opened, and the apparatus was exhausted by means of a Fleuss pump, to a pressure of 3 mm. of mercury. The tap s_3 was closed, t_3 opened momentarily, then reclosed, and the exhaustion was repeated. After a repetition of this washing process, the gas mixture from E was allowed to flow into the rest of the apparatus, t_3, t_4 and s_3 being closed, until the mercury height in A showed that the pressure throughout was nearly atmospheric, and approximately equal to that of the reserve gas in B. Finally t_3 and t_4 were opened; the pressure throughout rapidly attained equilibrium, t_2 was closed, and simultaneous readings were taken of the time and the heights of the liquids in D and B. t_1 was now closed, and the furnace was slowly rocked, at that rate which just allowed the catalytic pellets to roll from one end of R to the other between each reversal of position. This ensured a constant stirring of the gases in R—without this, the measurements recorded would be observations of gaseous diffusion, not of the rate of reaction.

While t_1 was closed, gas was slowly absorbed by the pressure tubing—this was counteracted by a raising of H sufficient to keep the water-level in D constant. After half-an-hour this absorption became negligible, and was always negligible in comparison with the absorption due to reaction between the gases at periods when the rate of combination was rapid. After a suitable interval, varying from one to ten minutes, the readings of the time and of B were noted, t_1 was opened and H rapidly adjusted to restore the constant pressure level in D. The new reading of B was observed, and t_1 was closed. These readings were repeated at intervals depending on the rate of reaction. During the "peak" portion of the reaction curve, the tap t_1 was left open all the time, as H had to be raised very rapidly and continuously. The reaction could not be followed beyond 120° C., the rate of absorption becoming too great to be measured by this apparatus.

Exhaustion of the reaction tube between successive experiments was the only revivification process employed. No diminution in catalytic activity was observed after three months of intermittent use.

Discussion of Results.

The curve showing change of reaction-velocity with time at constant temperature is of an unexpected form and presents some unusual features.

In most of our experiments there was a period, varying in duration from three hours to a few seconds after the mixed gases had been brought into contact with the nickel, during which time practically no reduction in volume, and therefore no hydrogenation, took place. This "induction" period was followed by a very rapid increase in reaction-velocity to a sharp maximum.

Series I.—Composition of Gas Mixture Constant, Temperature varying.

No. of expt.	II.		XIII.		XI.		I.		IX.		III.		XII.	
Temperature.	73° C.		81° C.		86·5° C.		98° C.		102° C.		103° C.		119° C.	
Per cent. C_2H_4.	50·5 per cent.		50·9 per cent.		50·9 per cent.		50·0 per cent.		50·0 per cent.		50·5 per cent.		50·7 per cent.	
	Time in mins.	Rate.	Time.	Rate.	Time.	Rate.	Time.	Rate.	Time.	Rate.	Time.	Rate.	Time.	Rate.
	Six hours.	No absorption observed.	0–141	No action	23	0·01	0– 52	No action	23	0·03	10	0·054	1½	1·31
			151	0·01	38	0·04	62	0·02	28	0·08	15	0·13	2	1·87
			163	0·06	44	0·08	72	0·03	33	0·69	20	1·14	3	2·71
			171	0·10	47	0·12	82	0·06	35	0·79	22	2·52	4	2·71
			175	0·09	51	0·32	87	0·26	36	1·33	24	1·11	5	2·15
			186	0·08	53	0·39	90	0·76	37	1·24	26	0·92	6	1·63
			207	0·08	57	0·64	92	1·16	39	0·95	28	0·92	7	1·78
			217	0·09	62	0·60	95	1·19	43	0·76	30	0·80	8	1·54
					69	0·55	97	1·15	46	0·67	33	0·67	9	1·45
					77	0·51	99	0·75	49	0·62	36	0·58	10	1·40
					83	0·48	101	0·64	53	0·57	39	0·52	12	0·94
					89	0·46	107	0·46	57	0·48	42	0·46	13	0·84
					95	0·39	111	0·38	61	0·43	46	0·42	15	0·72
					101	0·35	—	—	64	0·38	49	0·37	18	0·59
					105	0·32	—	—	70	0·31	52	0·35	21	0·53
									75	0·24	—	—	24	0·42
													27	0·42
													31	0·37

Note.—The "induction period" in Experiment XI is abnormally brief, probably owing to an incomplete removal of the reaction products of the preceding experiment.

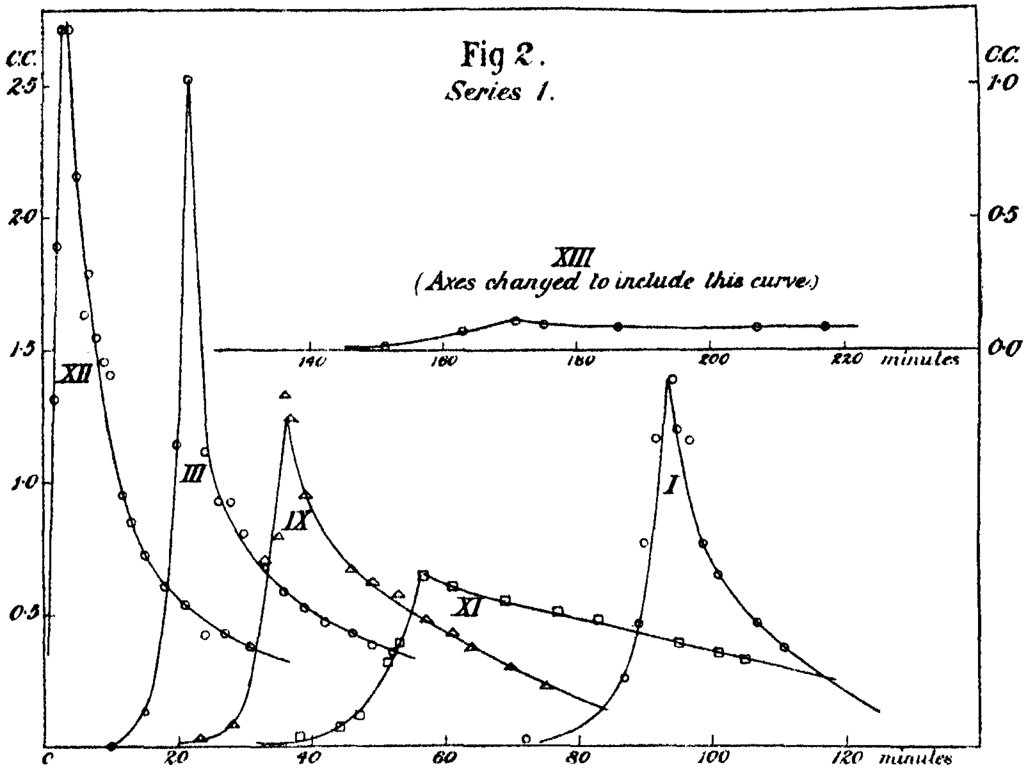


Fig 2.

Series 1.

Finally, the rate of reaction fell rapidly away from this maximum to a nearly steady value.

The initial period of inactivity might possibly be due to the selective adsorption of oxygen contaminating the electrolytic hydrogen. This oxygen would be gradually removed by combustion with hydrogen or ethylene.

Such an explanation would require the duration of the "induction" period at a given temperature to be approximately proportional to the hydrogen-content of the gas mixture, which an inspection of the series of results for different gas mixtures at 104° shows not to be the case:—

Temperature 104° C.

Duration of induction period in minutes.	Percentage by volume of hydrogen in mixture.
150	75
40	65
22	50
Very short	35
	25
	15

Series II.—Composition of Gas Mixture varying, Temperature nearly Constant.

No. of experiment.	IV.		X.		III.		VI.		VII.		VIII.	
Temperature.	103·5° C.		105·5° C.		103° C.		105° C.		103° C.		103° C.	
Per cent. C_2H_4.	25·5 per cent.		35 per cent.		50·5 per cent.		65·2 per cent.		75 per cent.		85 per cent.	
	Time in minutes.	Rate.	Time.	Rate.	Time.	Rate.	Time.	Rate.	Time.	Rate.	Time.	Rate.
	0–108	No action	11	0·048	10	0·054	1	2·57	1	0·56	2	0·07
	118	0·018	21	0·04	15	0·13	2	2·11	3½	1·09	5	0·09
	138	0·03	36	0·12	20	1·14	3	1·93	5	1·85	7	0·29
	148	0·13	40	0·71	22	2·52	5	1·20	6	1·02	9	0·35
	150	0·23	41	1·47	24	1·11	7	0·87	7	0·83	11	0·37
	151	0·42	42	0·95	26	0·92	8	0·74	10	0·65	13	0·30
	154	0·33	44	0·71	28	0·92	10	0·64	12	0·46	15	0·30
	155	0·28	45	0·62	30	0·80	11	0·55	15	0·37	17	0·28
	157	0·23	47	0·57	33	0·67	14	0·46	18	0·28	19	0·21
	160	0·24	49	0·50	36	0·58	17	0·37	20	0·23	20	0·19
	162	0·23	53	0·45	39	0·52	20	0·28	26	0·15	25	0·18
	165	0·24	57	0·39	42	0·46	22	0·23	33	0·08	30	0·093
	169	0·28	61	0·34	46	0·42	23	0·18	38	0·07	35	0·074
	172	0·28	67	0·29	49	0·37	28	0·15	—	—	41	0·056
	175	0·28	74	0·24	52	0·35	46	0·032	—	—	46	0·037
	178	0·28	82	0·19								
	181	0·28	96	0·12								
	185	0·25	106	0·09								
	188	0·22										
	192	0·22										
	196	0·19										
	201	0·15										
	213	0·13										
	233	0·08										

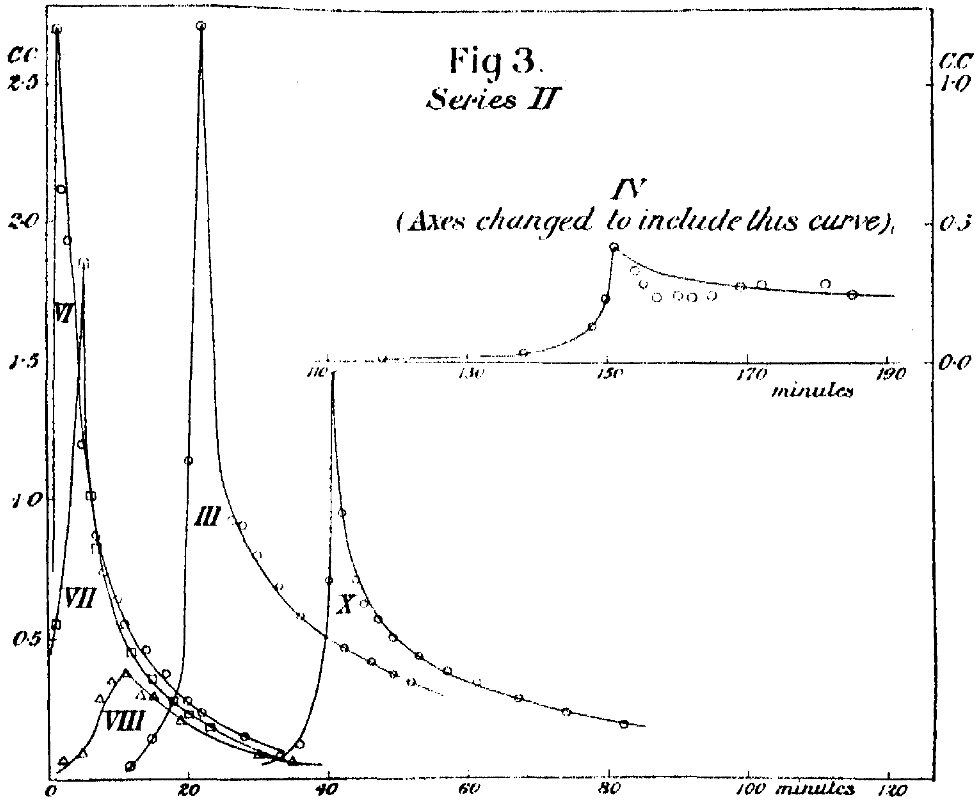


Fig 3.
Series II

A complete explanation of the form of the reaction-velocity-time curves can be put forward on the basis of the following assumptions:—

(*a*) Hydrogen is selectively adsorbed on nickel from mixtures of all the compositions studied.

(*b*) Ethylene is adsorbed only after activation (chiefly by thermal contact with a molecule undergoing hydrogenation).

On bringing the gas mixture into contact with the nickel, this becomes at once covered almost completely with hydrogen, only a few ethylene molecules being present in the adsorption layer; reaction with the formation of ethane takes place with these and the adsorbed hydrogen. The heat liberated in the hydrogenation is at once communicated to the molecules surrounding the hydrogenated molecule, leading to volatilisation of hydrogen and activation, followed by condensation, of ethylene.

The relative amounts of hydrogen and ethylene adhering to the surface of the catalyst thus change rapidly, once a number of centres of reaction have developed. This stage corresponds to the maximum in the reaction-velocity, Finally, the catalyst becomes covered almost completely with ethylene, to the exclusion of hydrogen.

Hydrogenation thus takes place comparatively slowly by collision between adsorbed ethylene and hydrogen or by diffusion of hydrogen to the catalyst surface through a thin layer of ethylene.

It will be noticed that the "induction" period is reduced in a very marked way either by increase of ethylene-content in the gas mixture or by increase in temperature, while the rate of reaction after the maximum has been reached is only very slightly affected by temperature. These observations are in accordance with the above theory. The final rate of reaction falls off gradually, owing to the accumulation in the reaction tube of inert ethane which acts as a diluent to the gas mixture.

The Catalytic Activity of Copper.—Part II.

By William George Palmer, St. John's College, Cambridge.

(Communicated by Sir Wm. Pope, F.R.S. Received April 14, 1921.)

In Part I of this series of papers,* the writer described the construction of an apparatus for determining directly the relation between reaction-velocity and temperature in the catalytic dehydrogenation of ethyl alcohol by the action of copper produced from cupric oxide by reduction. Some preliminary results were also given. This work has now been extended both by using reducing agents other than hydrogen (namely, carbon monoxide and methyl alcohol vapour) in the preparation of the catalyst and by the use of other alcohols as reacting substances.

In the earlier work, the method consisted in obtaining an automatic record of the rate of escape of hydrogen liberated in the reaction, at several different temperatures, the temperature being kept constant during each experiment. This method proved very laborious, since only by making experiments at a great number of different temperatures could a sufficiently intimate knowledge of the influence of temperature be obtained. Also, the catalyst had always to be reprepared by oxidation and reduction between experiments.

It has now been found possible to obtain a continuous record of the change of activity, while the temperature of the reaction is gradually increasing through the range to be studied.

---

* 'Roy. Soc. Proc.,' A, vol. 98, p. 13.

The gas-meter described in the former paper was not found suitable for the record of changing rates in the escape of hydrogen. After some tentative experiments, the following simple arrangements of recording apparatus were adopted. The hydrogen liberated in the reaction tube passed first into a flask of 1 litre capacity, kept surrounded with cold water, to secure temperature control of the gas, and thence to the gas-meter, which was very simply constructed of a glass tube, with a wide, salver-shaped end dipping into a wider vessel nearly filled with water.

With a head of water of 2 cm., the relation summarised below was found between the capacity of the meter and the rate of flow of gas.

Time in seconds for escape of 10 bubbles.	Size of bubbles in c.c.	Time in seconds for escape of 10 bubbles.	Size of bubbles in c.c.
12·8	2·18	70·0	1·86
17·2	2·05	93·0	1·84
56·8	1·87		

Thus the capacity of the meter is practically constant if a rate of flow of ten bubbles in one minute is not exceeded: the maximum rate of flow of hydrogen in any experiment was never greater than nine bubbles per minute.

The electrolytic manometer before described (*loc. cit.*) was again used, but not, in the present experiments, to actuate the chronograph directly. The electrodes were arranged as part of a telephone circuit in such a way that contact was made just before a bubble was released, and sharply broken at the release by the descent of the sulphuric acid. Immediately the "break" was heard in the telephone, a key was depressed by hand to actuate the chronograph, and, as nearly as possible at the same instant, the temperature was read on the millivoltmeter. Owing to the use of this telephone device, visual attention could be wholly concentrated on the accurate reading of the temperature. As the temperature reached its upper limit, it became necessary to record several bubbles as one on the chronograph to avoid having readings too close together, so that it was not possible to form a completely automatic record as formerly.

From the chronograph record two curves were constructed, one showing the relation between temperature and time, and the other that between reaction-velocity and time. It was then easy to make a simple reaction-velocity temperature curve from these two auxiliary curves. The exactness in coincidence on the two auxiliary curves of a temporary fall in temperature, with a corresponding fall in reaction-velocity, demonstrated the nicety with which the two curves were "in register."

The catalyst used in all the experiments was the same as that used in the work described before. At the close of the experiments, the copper was removed by solution in nitric acid, and estimated electrolytically as 0·022 grm.

A. *Reduction of the Oxide with Carbon Monoxide: Anhydrous Ethyl Alcohol as Reacting Substance.*

The carbon monoxide was prepared by passing carbon dioxide over strongly heated wood-charcoal. The resulting gas was passed through a long tube containing reduced copper at 300° C., to remove any possible contamination. The purified gas was stored in aspirators containing water, from which dissolved air had been removed by a current of carbon dioxide.

Experiments were conducted in a way almost exactly similar to those in which hydrogen was used as reducing agent. The apparatus was filled with pure dry nitrogen instead of with carbon dioxide as formerly.

The necessary increase of sensitiveness in the millivoltmeter was obtained by using a bath of boiling ethyl benzoate (boiling point 213° C.) for the constant temperature junction, and an external resistance was dispensed with. An auxiliary couple in a boiling water bath was used to determine reduction temperatures below 213° C. The period of heating from about 230° C. to 300° C. was usually about one and a quarter hour. The reduction of the copper oxide was always signalised by a "kick" on the millivoltmeter needle over about 4° C. The mean between the highest temperature reached and the initial temperature was taken as the temperature of reduction.

Table I.—Catalyst prepared by reduction with Carbon Monoxide: Anhydrous Ethyl Alcohol as Reacting Substance.

(1) Temperature of Reduction 215° C. (Curve III, fig. 1.)

Temperature.	Volume of hydrogen in 30 seconds.	Gramme-molecules of alcohol decomposed per second per gramme-atom of copper × 10^3.	
		Observed.	Calculated.
° abs.	c.c.		
499	1·1	4·77	3·82
507	1·5	6·50	6·30
514	2·4	10·4	10·8
519	4·0	17·3	15·9
525	5·6	24·3	24·1
529	7·2	31·2	31·2
533	8·8	38·2	42·0
537	10·3	44·6	44·6
538	10·9	47·3	
539	11·0	47·7	
541	11·1		
545	11·3		
550	11·6		
559	12·5		

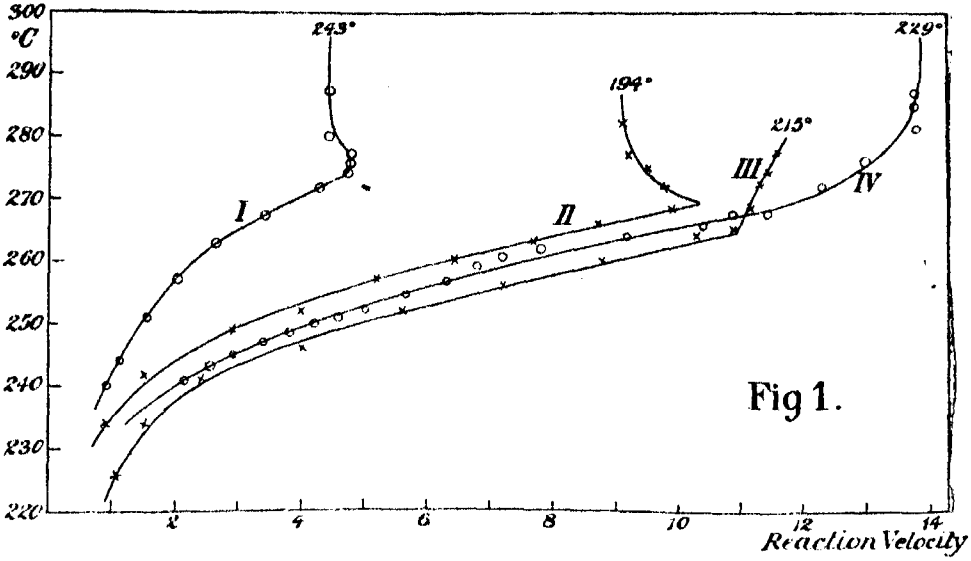


Fig 1.

(2) Temperature of Reduction 243° C. (Curve I, fig. 1.)

Temperature.	Volume of hydrogen in 30 seconds.	Gramme-molecules of alcohol decomposed per second per gramme-atom of copper × 10^3.	
° abs.	c.c.	Observed.	Calculated.
513	0·9	4·32	3·7
517	1·1	4·77	4·8
524	1·5	6·50	6·5
530	2·0	8·66	8·7
536	2·6	11·20	12·0
540	3·4	14·70	14·4
545	4·25	18·4	18·4
548	4·7	20·4	20·5
549	4·7		
550	4·75		
553	4·75		
560	4·4		

(3) Temperature of Reduction 229° C. (Curve IV, fig. 1.)

Temperature.	Volume of hydrogen in 30 seconds.	Gramme-molecules of alcohol decomposed per second per gramme-atom of copper × 10^3.	
		Observed.	Calculated.
° abs.	c.c.		
514	2·1	9·10	9·0
518	2·9	12·6	12·6
520	3·4	14·7	14·8
523	4·2	18·2	18·1
528	5·7	24·7	24·7
530	6·3	27·3	29·5
534	7·2	31·2	
537	9·2	39·9	
539	10·4	45·1	52·6
545	12·3		
549	13·8		
554	13·8		

(4) Temperature of Reduction 194° C. (Curve II, fig. 1.)

Temperature.	Volume of hydrogen in 30 seconds.	Gramme-molecules of alcohol decomposed per second per gramme-atom of copper × 10^3.	
		Observed.	Calculated.
° abs.	c.c.		
498	0·5	2·16	2·16
500	0·6	2·60	2·57
507	0·9	3·90	4·31
512	1·4	6·07	6·07
517	2·1	9·10	9·00
522	2·9	12·6	12·6
525	4·0	17·3	16·1
530	5·2	22·5	22·4
533	6·4	27·7	27·7
536	7·7	33·4	33·9
539	9·2	39·9	41·0
542	10·3	44·6	
545	9·8		
548	9·5		
550	9·2		
554	9·1		

(5) Temperature of Reduction 156° C. (Not shown in figure.)

Temperature.	Volume of hydrogen in 30 seconds.	Temperature.	Volume of hydrogen in 30 seconds.
° abs.	c.c.	° abs.	c.c.
502	1·0	532	7·5
509	1·1	534	8·2
515	1·5	542	9·4
520	2·8	545	9·0
525	4·3	550	9·3
529	5·9	561	10·0

B. *Reduction of the Copper Oxide with Methyl Alcohol Vapour.*

A small distilling flask was sealed on to the top of the reaction tube in the position of the delivery tube for carbon monoxide or hydrogen. Pure anhydrous methyl alcohol was drawn into the flask when the reduction temperature in the reaction tube was reached. A slow current of nitrogen was passed through the flask and the methyl alcohol gently boiled by an external heating coil. On opening the lowermost tap of the apparatus the stream of vapour was carried through the reaction tube over the oxide. The minimum temperature of reduction for methyl alcohol vapour was found to be 205° C.

The catalytic effects of the catalyst prepared in this way are summarised in Table II:—

Table II.—Catalyst prepared by Reduction with Methyl Alcohol Vapour: Anhydrous Ethyl Alcohol as reacting substance.

(1) Temperature of Reduction 217° C. (Curve X, fig. 2.)

Temperature.	Volume of hydrogen in 30 seconds.	Temperature.	Volume of hydrogen in 30 seconds.
° abs.	c.c.	° abs.	c.c.
501	0·9	527	6·8
504·5	1·2	532	7·2
511	2·35	534	7·4
518	4·1	539	7·9
522	5·8	542	8·9

(2) Temperature of Reduction 220° C. (Curve IX, fig. 2.)

Temperature.	Volume of hydrogen in 30 seconds.	Temperature.	Volume of hydrogen in 30 seconds.
° abs.	c.c.	° abs.	c.c.
522	3·0	540	7·0
524	3·6	544	8·2
527	3·8	546	8·5
534	5·6	549	8·6
537	6·4		

(3) Temperature of Reduction 240° C. (Curve VIII, fig. 2.)

Temperature.	Volume of hydrogen in 30 seconds.	Temperature.	Volume of hydrogen in 30 seconds.
° abs.	c.c.	° abs.	c.c.
518	0·70	540	6·70
522	1·50	544	7·6
527	2·50	547	7·9
530	3·65	549	7·9
534	4·75	552	7·9
537	5·75		

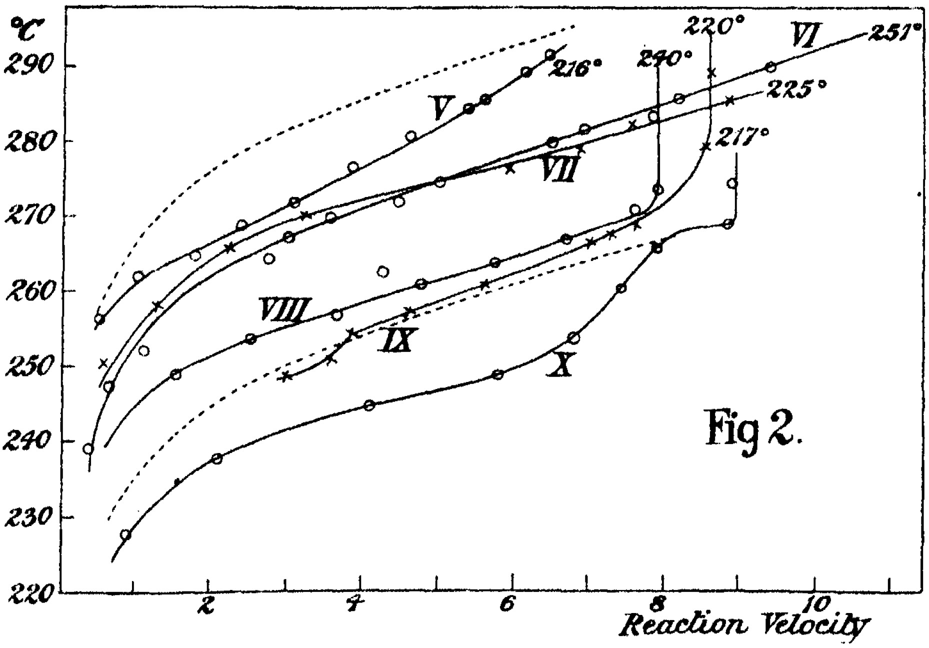


Fig 2.

C. *Experiments with Isopropyl Alcohol as the Reacting Substance.*

The alcohol used was a mixture of an old stock made from acetone by Sabatier's method with a new sample made by reducing acetone with sodium; this sample was supplied to me by the kindness of Mr. H. A. Scarborough.

The mixed samples were fractionated to within 0·1° C. by using a Dufton column. The constants of the alcohol then showed very satisfactory agreement with the established values for the constant-boiling mixture of *iso*propyl alcohol and water. Owing to the well-known difficulty of removing the last traces of water from this alcohol, its dehydration was not attempted. It was considered more satisfactory to ensure that all the vapour passing through the reaction tube should be of constant composition than to use a partially dehydrated alcohol, giving a vapour of variable composition.

Owing to lack of time, the hydrogenation of this alcohol was studied only with the catalyst prepared with methyl alcohol vapour. The results show a general agreement with those for ethyl alcohol when the catalyst had been prepared with methyl alcohol or carbon monoxide.

Table III.—Catalyst prepared by reduction with Methyl Alcohol Vapour: Isopropyl Alcohol with 12 per cent. of Water as Reacting Substance.

(1) Temperature of Reduction 225° C. (Curve VII, fig. 2.)

Temperature.	Volume of hydrogen per 30 seconds.	Gramme-molecules of alcohol decomposed per second per gramme-atom of copper × 10^3.	
° abs.	c.c.	Observed.	Calculated.
513	0·5	2·17	2·0
520·5	0·65	2·82	3·1
525	1·1	4·77	4·6
531·5	1·35	5·85	6·0
537·5	2·75	11·9	11·1
540	3·2	13·9	13·5
543	3·55	15·4	15·4
545	4·45	19·3	19·0
553	6·5	28·2	27·6
557·5	6·9	29·9	30·0

(2) Temperature of Reduction 216° C. (Curve V, fig. 2.)

Temperature.	Volume of hydrogen per 30 seconds.	Temperature.	Volume of hydrogen per 30 seconds.
° abs.	c.c.	° abs.	c.c.
529·5	0·5	554·5	4·60
538	1·75	558	5·35
541	2·35	561	5·90
545	3·05	563	6·10
550	3·80	566	6·60

(3) Temperature of Reduction 251° C. (Curve VI, fig. 2.)

Temperature.	Volume of hydrogen in 30 seconds.	Temperature.	Volume of hydrogen in 30 seconds.
° abs.	c.c.	° abs.	c.c.
522	0·90	551	5·0
526	1·10	556	6·4
531·7	2·00	559	7·0
538	2·8	563·5	7·9
542·2	3·8	566	8·4
547	4·3	573	9·6

These results are shown graphically in figs. 1 and 2.

Discussion of the Results.

1. It will be seen that the activity of the catalyst does not necessarily increase continuously as the temperature of preparation from the oxide is lowered.

The order of activity is as follows:—

Temperature of reduction ° C.

(Carbon monoxide)
fig. 1.
215
229 Ethyl alcohol as reacting substance.
194
243 (The curve for preparation at 156° will be discussed later.)

(Methyl alcohol
vapour) fig. 2.
251
225 (Anomalous above 270°.)
217 Isopropyl alcohol as reacting substance.

This result conflicts with the view that the real catalytic agent is an oxide of copper dissolved in reduced (metallic) copper, but it could be readily reconciled with the theory put forward in the previous paper; namely, that metallic copper produced by reduction of cuprous oxide is the active agent. The proportion of active material formed by reduction would depend on the mutual relation of the velocity curves for the two reactions

$$CuO + H_2 = Cu + H_2O ; \qquad 2CuO + H_2 = Cu_2O + H_2O.$$

In support of this hypothesis, it may be mentioned that copper reduced from cuprous oxide is more readily attacked by ordinary reagents than metal produced from cupric oxide.*

It will appear from the later discussion that the inactivity of "cupric" copper is to be attributed to its lack of adsorbing power for alcohol, and not to any other cause.

2. The curves numbered in the figures 1 and 2, I, II, III, IV, VI, VIII, VII (below 270°), strictly conform to the relation $v = Ae^{-K/T}$ up to the point where the sudden change of direction occurs. Values of v calculated from this law are shown in the last column of the above Tables.

$\log_e v$ was plotted against $1/T$, and from the resulting straight lines the following values were obtained:—

---

* Private communication from Mr. E. K. Rideal.

Table IV.

Reducing agent, carbon monoxide : ethyl alcohol as reacting substance.

Reduction temperature.	K.	$\log_e$ A.	Activity coefficients.	v 260° C. ÷ v 250° C. (from $\log_e v$ values).
° C.				
215	19,380	46·90	$2·4 \times 10^{-3}$	2·00
229	19,360	46·76	$2·1 \times 10^{-3}$	2·00
194	19,390	46·5	$1·6 \times 10^{-3}$	2·00
243	14,230	35·9	$4·0 \times 10^{-8}$	1·67

Reducing agent, methyl alcohol vapour : isopropyl alcohol as reacting substance.

251	19,490		$4·8 \times 10^{-4}$	2·01

The simplest way of interpreting this temperature-velocity law would consist in imagining the catalytic process to resemble the thermionic emission of electrons from hot metals.

The active surface of the catalyst is to be regarded as covered with a layer of alcohol several molecules deep. Decomposition of alcohol, however, only occurs (to any measurable extent) in the layer in immediate contact with the catalyst. "Activation," that is, the absorption of energy necessary before reaction occurs, is brought about by absorption of radiation from the atoms of the catalyst. It is assumed that only in the layer of alcohol molecules next to the catalyst is the intensity of this radiation sufficient to bring about activation. This view receives support from the effect of adsorbed hydrogen on the catalyst to be discussed at a later point. Also, it appears to harmonise the hitherto opposed views of catalysis—the "physical" adsorption theory which made no allowance for the specific action of catalysts, and the "chemical compound" theory against which much has been urged. On the theory advanced above the specific action of the catalyst lies not only in selective adsorption but in providing radiation of the proper quality. Langmuir* has indicated that collision between gaseous molecules and a solid differs from ordinary kinetic collision in a homogeneous gas phase in being comparatively inelastic. The colliding gas molecule therefore sojourns on the surface of the catalyst for a period long compared with that occupied in ordinary "gaseous" collision and increased opportunity is given for activation. Herein lies the essential difference between catalysed and uncatalysed reactions.

* 'Phys. Rev.,' vol. 8, p. 149 (1916).

On the basis of these views a very simple explanation can be given of the sudden change in direction of the curves at temperatures approximately between 265° and 275° C.

While the alcohol film remains several molecules deep, catalytic action is independent of the rate of effective condensation on the catalyst, since the ranks of molecules always remain closed. As the temperature is increased the thickness of the adsorbed layer diminishes owing to the increased rate of evaporation, and at about 270° C. the film is only about two molecular diameters or less in thickness.

The rate of decomposition will now become dependent on an additional factor—the time-difference between condensation and evaporation of an alcohol molecule. Practically nothing is known concerning this time-difference or "life" of a molecule on a solid surface, but the effect on the catalytic activity of its variation with temperature might be indicated in a hypothetical extreme case by supposing the "life" of a molecule to be given by

$$\pi = \mathrm{B}e^{a\mathrm{T}},$$

where B and a are constants.

Then the catalytic activity is given by

$$v = \mathrm{AB}e^{-\mathrm{K}/\mathrm{T}} \,.\, e^{a\mathrm{T}} = \mathrm{AB}e^{-a\mathrm{K}},$$

that is, the catalytic activity is independent of the temperature. The curves indicate a very great decrease of the temperature coefficient and in some cases (as in curves I, II, XI, XII, figs. 1 and 3) even a decrease of the actual catalytic activity, due to the introduction of a new factor as indicated.

In column 4 of Table IV are shown a series of numbers called activity coefficients, calculated in the following way. The maximum number of alcohol molecules that could be condensed in a layer one molecule deep on one gramme-atom of copper would be the Avogadro number $\mathrm{N} = 6{\cdot}06 \times 10^{23}$, assuming all the copper atoms could each adsorb one molecule of alcohol. A measure of the maximum number of alcohol molecules activated by the catalyst is given by the value of A in the equation

$$v = \mathrm{A}e^{-\mathrm{K}/\mathrm{T}}.$$

The activity coefficients show the value of the ratio A/N, which measures the combined effect of number of "active" copper atoms and the number of effective condensations.

The curves for *iso*propyl alcohol show no sudden change in direction up to 300° C.; presumably, the layer of this alcohol remains more than two molecular diameters thick above this temperature.

3. If the quantum theory of radiation is adopted, we may write

$$\frac{\mathrm{N}h\nu}{\mathrm{R}} = \mathrm{K},$$

where the symbols have their usual significance, and K is given by

$$v = \mathrm{A}e^{-\mathrm{K}/\mathrm{T}}.$$

For the curves II, III, IV, VI (figs. 1 and 2)

$$\nu = 4{\cdot}04 \times 10^{14}, \qquad \lambda = 0{\cdot}742\,\mu.$$

For curve I (fig. 1)

$$\nu = 3{\cdot}00 \times 10^{14}, \qquad \lambda = 1{\cdot}0\,\mu.$$

The heat of activation of alcohol in calories is given by

$$\mathrm{RK} = 1{\cdot}98 \times 19{,}380 = 38{,}740 \text{ calories per gramme molecule.}$$

This value may be used to calculate the heat of activation of the reversed reaction, aldehyde + hydrogen → alcohol, by using Haber's relation

$$\Sigma \mathrm{N}h\nu \text{ (reactants)} - \Sigma \mathrm{N}h\nu \text{ (resultants)} = \text{heat of reaction.}$$

The heat of reaction may be found in the usual way:—

$$[\mathrm{CH_3CHO,H_2}] \longrightarrow \mathrm{CO_2 + H_2O} + 347 \text{ kgrm. calories.}$$

$$[\mathrm{CH_3CH_2OH}] \longrightarrow \mathrm{CO_2 + H_2O} + 326 \text{ kgrm. calories.}$$

giving

$$\mathrm{CH_3CHO + H_2 = CH_3CH_2OH} + 21 \text{ kgrm. calories,}$$

whence the total heat of activation of the system [aldehyde + hydrogen] is 17 kgrm. calories, or approximately half the value for alcohol. Hence the hydrogenation of aldehyde requires λ = approximately $1{\cdot}6\,\mu$.

The infra-red spectrum of copper shows absorption bands at $\lambda = 0{\cdot}8\,\mu$ and at $\lambda = 1{\cdot}6\,\mu$, so that the catalyst could provide the proper frequencies for both reactions, the lower frequency being more strongly developed at lower temperatures.

These considerations suggest an explanation of the commonly observed fact that metallic catalysts can operate in either direction of a reversible reaction according to the temperature.

The Effect of "Foreign" Substances Adsorbed on the Catalyst.

The curves numbered V, X in fig. 2 depart markedly, and VII and IX slightly, from the simple logarithmic form, in the direction of increased activity between certain temperatures.

This abnormality is most marked with catalyst material prepared by a reducing agent that gives water as a product of reduction, namely, hydrogen and methyl alcohol, and increase in the effect is also shown by reduction at comparatively low temperatures.

It was thought that water might be retained by the catalyst in these cases and exert a strong specific catalytic action. To test this supposition, a number of curves were obtained for constant-boiling mixtures of various alcohols with water. These are shown chiefly in fig. 3. The broken line curves indicate the course of the simple logarithmic law. It will be noticed (in fig. 2) that, when *anhydrous* ethyl alcohol is used, a catalyst prepared with methyl alcohol at a temperature above 220° does not retain water, but, in spite of this, curves XI and XII show that water can be adsorbed on the catalyst from the reaction mixture (roughly in proportion to the amount present) even when the catalyst has been prepared at temperatures of 232° C. and 239° C. When isopropyl alcohol with 12 per cent. of water is used as reacting substance, water is apparently not retained on the catalyst if this is prepared above a temperature of about 216° (curve V), although curve VII shows a very slight deviation from the law, which is strictly obeyed by curve VI. The following is a list of the aqueous reagents, with the appropriate curves:—

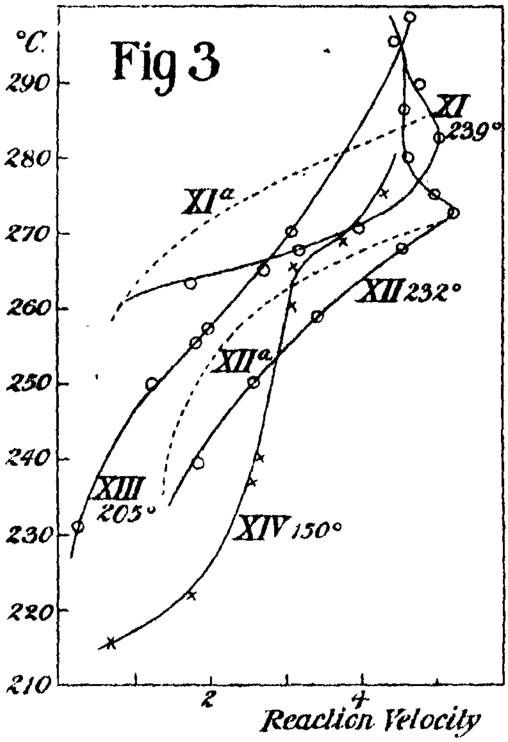


Fig 3

Ethyl alcohol + 4·5 per cent. water (curve XII, fig. 3).
Normal propyl alcohol + 28 per cent. water (curve XI, fig. 3).
Isopropyl alcohol + 12 per cent. water (curves V, VI, VII, fig. 2).

This "promoter" effect of water could be predicted on the radiation theory, since the chemical similarity of water and the aliphatic alcohols would lead to similar spectra in the infra-red region.

The curve for reduction by carbon monoxide at a temperature of 156° (not given in figure) shows a slight departure from the logarithmic law, presumably due to traces of water in the ethyl alcohol.

The two curves for hydrogen reductions (*loc. cit.*) are included in order to discuss them in the present connection (curves XIII, XIV, fig. 3).

The low activity of the catalyst prepared by hydrogen must be attributed to the well-known retention of hydrogen by the reduced copper, which is consequently "poisoned." The catalyst is not completely covered with hydrogen, as the "water-effect" is very marked on both curves. Hydrogen has absorption bands in the infra-red at 0·6 μ and at 1·2 μ, so that it would not be expected to act as a catalytic agent in the dehydrogenation, nor apparently is the radiation from the underlying copper able to act to any considerable extent on alcohol molecules across the hydrogen layer.

The work is being continued, with a view to finding the effect of admixed metals and oxides on the catalyst.

A Reciprocating Expansion Apparatus for Detecting Ionising Rays.

By Takeo Shimizu.

(Communicated by C. T. R. Wilson, F.R.S. Received June 4, 1921.)

The present investigation was undertaken for the purpose of obtaining a very sensitive method of detecting ionising rays, such as X-rays and rays from radio-active substances. An electroscope or electrometer, in combination with an ionisation chamber, even if the required sensitiveness were obtained, would not distinguish between the effects of a single ray and of two or more which entered the chamber simultaneously. In order to attack certain problems relating to the structure of ethereal waves and other delicate questions, it is necessary to have some means of detecting and discriminating individual rays. Mr. C. T. R. Wilson's beautiful method of making visible the tracks of ionising rays in gases[*] affords us the most powerful means for that purpose, and the writer accordingly tried to obtain a modified form of the original apparatus, such as could be readily used for various experiments.

---

[*] 'Roy. Soc. Proc.,' A, vol. 87, pp. 277–292 (1912).

One important quality such apparatus should possess is the ability to correlate the number of tracks which are registered by it with the time interval during which the rays which produce them are emitted. With the original method, the expansion takes place so suddenly that it is difficult to determine this interval with accuracy, although it must be remembered that such suddenness has the merit of yielding well defined tracks at the moment of expansion. Another point is that the possibility of registering tracks, by producing the appropriate supersaturation of the gas in the cloud chamber, should be available continuously for the few hours which is usually required for any series of observations.

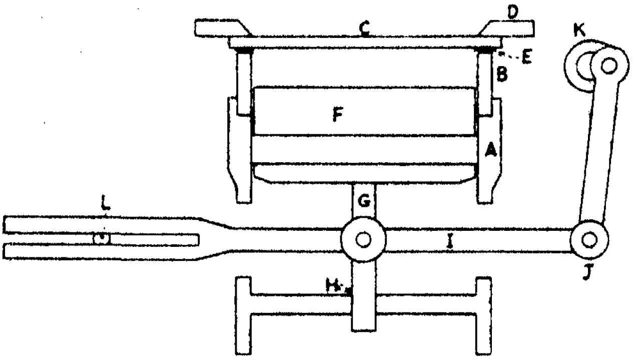
As it seemed difficult to provide a means to keep a volume of air constantly at a high supersaturation (which would be the most satisfactory way if it were possible), attempts were made to comply with the above conditions by using the original method of expansion, repeated at a convenient frequency, and also at a moderately slow rate of expansion.

There exist a number of imaginable mechanical and physical difficulties even for this simple method. It was doubted whether such a slow expansion would be good enough for the formation of a water drop around an ion. It was also expected that the air might become too dry after a number of expansions, the drops carrying much of the moisture to the bottom, and that such repetition of expansion might create a turbulent motion of air and distort the individual tracks. But, after some trials, it was found that the difficulties of the method were not so serious as had been feared.

The actual apparatus so far constructed is very similar to that of the original method, the chief difference being only in the mode of working the plunger or the piston. Instead of adjusting the position of the plunger before each expansion, according to the expansion ratio desired, and then letting the air pressure on the plunger suddenly press it down, a mechanical connection is made from a prime mover to the piston, so that the latter makes a reciprocating motion between two definite positions. The nature of the motion has much to do with the distinctness of the tracks, but, so far, an approximately harmonic motion has been used, because of its mechanical simplicity. The frequency of reciprocation can be anything from one to five or more per second, but the best number for use seems to be about two per second. Another feature with the present apparatus is a commutator, which establishes and destroys an electrostatic field between the upper and the lower surfaces of the chamber in synchronism with the piston.

The main parts are sketched in the figure.

A is a brass cylinder of suitable size fixed in position. B is a glass cylinder of the same inner diameter as A, and is fixed in A so that the inner walls of the two make one continuous surface. C is a circular glass plate, which forms the upper boundary of the chamber, and is firmly pressed down by a brass ring, D, large enough not to obstruct the view through C. E is a thin copper ring with an electric lead. All the surfaces of contact of the various parts of the chamber are made air-tight by putting thin rubber sheets cut into ring shape between, and applying a pressure through D by means of set screws. F is a brass piston of sufficient length, with an effective leather packing at its lower end. The proportions are such that the upper surface of the piston reciprocates within the lower half of the glass cylinder, while the packing always remains well inside the brass cylinder. The body of the piston must be as closely fitting as possible to the cylinder, but must not touch the inner surface of the glass walls at any point. It is important that the sliding surface between the packing and the cylinder be air-tight all through the stroke, and it was found satisfactory to use tap grease as lubricant, care being taken that it did not contaminate the glass walls. The piston is connected to a rod, G, which is guided by a stationary hole, H, in the base of the apparatus. A horizontal rod, I, is passed through G near the centre of its length and pinned to it. One end, J, of this rod is connected to a horizontal shaft, K, which is driven by a motor, by means of an arm and a link; J thus executes a rocking motion when the shaft rotates. At the other end it is split into two, and is penetrated by a horizontal stationary pin, L, which serves as a fulcrum. This pin, although fixed in position when the apparatus is properly running, is capable of adjustment in a horizontal line passing through the axis of the cylinder. The dimensions of all other parts being fixed, the position of this pin will determine the amount of stroke of the piston, that is, the ratio of expansion of the air, and it is convenient to be able to adjust this ratio while the apparatus is running. Although a moderately wide range of the ratio gives



equally good results, and the ratio which is supposed to be the best on one occasion will slightly differ from that on another according to the temperature of the room, the humidity in the chamber, the frequency of the reciprocation, and so on, 1·7 may be taken as not far from the best mean ratio. Too small an expansion often leaves the tracks incomplete or diffuse, while a too large one spoils them by the formation of a general cloud.

The mean height of the air space in the chamber has also a great deal to do with the result, and it is better to make it small, for the reasons that, with a shallow chamber, water drops will settle down to the bottom quicker, and the air will be less subject to disturbing motion, other conditions being the same. About 1 cm. seems to be a convenient height.

The upper surface of the piston is a little recessed from its edge, and is filled with blackened moist gelatine (about 15 grm. of dry gelatine and a little Indian ink in 100 c.c. of water), and the lower surface of the glass roof of the cloud chamber is coated with well conducting moist gelatine (about 10 grm. of gelatine and 2 grm. of copper sulphate in 100 c.c. of water), care being taken to let the copper ring (E) and the gelatine film make good contact with each other all round the circumference. It is better to give the gelatine films a moderate thickness (1 mm. or more), to avoid quick drying when the cover is temporarily removed for other purposes. When the gelatine surfaces either collect some dust particles or become too dry, the surfaces may be wiped with moistened soft cotton. To make sure that the transparent lining on the cover is of uniform thickness so as not to distort the image of the tracks, it is convenient to look, while the gelatine is still fluid, at the two images of a distant bright source reflected from the surfaces of the gelatine and of the glass, and adjust the level so that the two images coincide. The cylindrical surface of the chamber is simply cleaned and left uncoated.

The shaft (K) is provided with a disc-shaped commutator (not indicated in the figure), composed of two insulated sectors, one permanently earthed, and the other kept at a negative potential of 200 volts or more. A stationary brush, which is connected to the lead of the copper ring (E), makes contact with those sectors alternately. In order to be able to adjust the relative durations of these contacts, while the commutator is rotating, a curved line is used instead of a straight radial line for one of the boundaries between the two sectors. The frame of the apparatus, and, in consequence, the piston, are permanently earthed. The phase of the commutator is adjusted so that the air space in the chamber is subjected to an electrostatic field, commencing at the beginning of the upward stroke, and ending at a certain stage of the following downward stroke, when the air acquires the supersaturation just sufficient to form a cloud on ions. This stage is arrived at when the expansion is about half completed, although of course it depends upon the amount of stroke and other conditions.

If no electric field is applied at all, rays which entered into the chamber prior to this stage will appear as diffuse bunches, instead of sharp lines, owing to the spontaneous diffusion of ions, the diffuseness depending upon the age of each track at the moment of critical supersaturation. These diffuse clouds are not only unsuitable for observation but hinder also the growth of younger tracks, as they absorb the vapour about them and reduce the supersaturation. The electric field serves to remove all the ions generated between two successive supersaturations, as well as those remaining since the previous expansion. During a supersaturation the chamber must be free from electric field, because the ions on a track of ray will be pulled apart from their original positions along the direction of the field, according to the sign of their charge, before they lose their mobility by accumulating a sufficient amount of water. The distance they travel in this way is not the same for all ions in a given field, but seems to be subject to a certain continuous distribution law. Hence the appearance of an α-ray track is not a line but a vertical ill-defined sheet of cloud, densest along a middle line where the ray passed originally. Thus it is important to destroy the field at the right moment, and this is done by adjusting the brush while the expansions are going.

When the apparatus is ready for use, it is convenient to test it by putting a small piece of metal sheet with a trace of polonium upon it on the inner surface of the glass cylinder. After the cover has been tightly screwed down the room is darkened and the chamber illuminated with a horizontal parallel beam of light. For eye observation a half-watt lamp with a condenser affords a convenient source. For photographing, a carbon arc or some other intense source is necessary. When the source produces a noticeable amount of heat, care must be taken to cool the beam by inserting a water trough before it reaches the chamber. Two diaphragms are used with advantage to cut off divergent rays so that the beam nowhere touches the two opposite horizontal surfaces of the chamber. It is the upper half of the cloud chamber that produces the best defined tracks, so that both the ionizing rays and the beam of light may be confined to that part only.

When the chamber contains fresh atmospheric air the strokes of the piston will simply make the water vapour collect upon dust particles floating in it and a dense uniform cloud will be seen at each expansion. But this will subside after a small number of expansions, the central portion of the chamber being the first to become free from cloud. If there be any slight leakage of air into the chamber it will at once make its appearance as a stream of cloud which indicates the position of the leak; this can generally be easily rectified. When there is leakage of electricity across the inner surface of the side walls as a result of condensed moisture bridging the two opposite surfaces, a dense fog will also appear; this can be distinguished from the air leakage without difficulty.

Reciprocation of the piston always creates a slowly moving vortex ring in the air, co-axial with the cylindrical wall and situated close to it. The sense of rotation is such that the air next to the wall travels upward. The main body of the air including the centre is practically free from motion of this kind, and the tracks keep their form for several expansions, appearing and disappearing as the reciprocation goes on, unless they are speedily swept away by a sufficiently strong electric field.

It is likely that some of the drops coalesce with each other when the track is a dense one, like that of an α-ray, and the drops are allowed to grow large enough. The result is that the free surface is decreased and the evaporation much retarded, and the tracks are apt to survive many expansions in spite of the intermittently applied electric field. This is disadvantageous because new rays passing near them cannot collect sufficient water, and often appear as discontinuous tracks, or look as if they started at a considerable distance from the source. A weak source of rays is always advantageous in this relation, for the chance of two rays passing the same region one immediately after another is small.

It is perhaps worth mentioning here some of the peculiarities observed about the cloud tracks. If the electric field is not cut off during the period of condensation, the track of an α-ray is spead out into a vertical sheet as stated above. And if the field is of the right magnitude a sharply defined clear space can be seen across the sheet along the original direction of the ray. The contrast in sharpness between its boundaries and the upper and the lower edges of the sheet is remarkable, which perhaps tends to show that there is a definite maximum in the size of the drops when first formed.

Very intense and continuous electric fields were once tried, in order to see if it were possible entirely to stop the appearance of tracks by rapidly pulling the ions to the walls before they could grow up into drops. This was done to see if it might be possible to limit sharply by means of the commutator alone the periods during which cloud tracks were formed. This stage was arrived at with a field of the order of 1,000 volts per centimetre, but at the same time the number of neutral water drops, which usually exist to a certain extent throughout the chamber, increased considerably, so as to form a general cloud, and the rays left narrow vertical clear spaces along their tracks. Thus a kind of reversal of effect takes place when the field is increased, a vertical sheet of cloud in clear space being replaced by a vertical clear space in a cloud.

The above method having proved to be an inconvenient one for confining the production of cloud tracks to a definite short interval during each cycle, a small mechanical shutter, placed in the chamber and worked from outside automatically with a small magnet, which exposes the source for a predetermined short interval during the supersaturation, was used for the investigation of α-rays. More penetrating rays were flashed through a window in the wall from outside.

For the cloud-forming substance other liquids than water, such as alcohol and petroleum, were also tried. They give about the same results as water, so far as the general appearance of the tracks is concerned.

Conclusion.

There is no doubt that Mr. C. T. R. Wilson's original apparatus is superior for obtaining the most accurate shape of each track. It is mainly as a detector that the present one is likely to come into use. If the circumstances are favourable to its use, an apparatus of this type can hardly be surpassed by any other instrument in sensibility, for the simple reason that it can detect the elementary charge. The other relative merit that this type possesses over the electroscope-ionisation chamber system is that with the latter the natural ionisation in the chamber always puts a limit to the practical sensibility, while with the former one can differentiate to a large extent the genuine from the stray rays by their shape and size. For instance, an α-ray track observed during an ordinary X-ray measurement can be left out of account.

In conclusion, the writer wishes to express his sincere gratitude to Sir E. Rutherford, in whose laboratory the present work has been carried out, for his kind encouragement and advice. His cordial thanks are also due to Mr. C. T. R. Wilson, whose keen interest and advice made the work possible, for his unfailing kindness.

A Preliminary Note on Branched α-Ray Tracks.

By TAKEO SHIMIZU.

(Communicated by Sir E. Rutherford, F.R.S. Received June 4, 1921.)

According to Prof. Sir E. Rutherford,* if about one hundred thousand α-rays from radium C pass through the air, on the average there will be one close nuclear collision, which produces a swiftly moving atom. Thus, if we take a great number of photographs of α-ray tracks compatible with the above order of magnitude, we might expect to find some evidence to indicate the disruption of atoms by the α-particles. The present note is to describe some preliminary trials in this direction.

The reciprocating expansion apparatus of the preceding article is very convenient for taking a large number of photographs within a reasonable time. Direct eye-observation confirmed the existence of some branched tracks which differed greatly in configuration from spurred tracks like the one photographed by Mr. C. T. R. Wilson.† Arrangements were then made to devise a suitable method of photographing such tracks and to show their orientation in space.

The use of cinematographic film seemed to be imperative and the camera was designed accordingly. It is run by the same shaft as the piston of the expansion chamber, so that the exposures are made automatically and synchronously with the apparatus. An ordinary photographic shutter is used and photographs are taken once in each expansion when the piston is at its lowest position, and therefore the air in the chamber is practically motionless. The time of exposure can be extended as long as one-fiftieth of a second without impairing the definition. The source of light employed is a carbon arc in combination with a condenser and a water-cooling device. The parallel beam of light, after passing through the chamber, is reflected back again by means of a plane mirror fixed on the other side of the chamber. The lens used is an Aldis F/3.

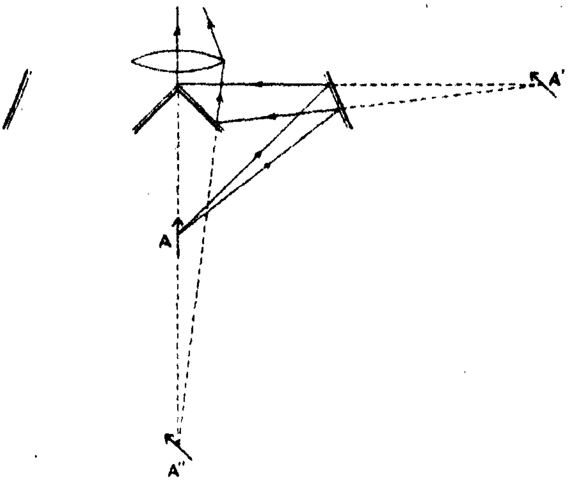
The special feature of this camera lies in the use of a mirror system which projects simultaneously two images of the object, viewed from mutually perpendicular directions, through the same objective lens. The principle is sketched in the figure. It is nothing but a variety of an ordinary range-finder.

Four mirrors with surface-silvering (indicated by a fine line in the figure) are used, two of which are fixed immediately in front of the lens, each making 45° with the vertical, and the other two at equal distances from the axis of the lens and at about the same level as the first pair, but inclined

---

* 'Phil. Mag.,' vol. 37, June, 1919; "Bakerian Lecture," 'Roy. Soc. Proc.,' A, vol. 97, 1920).

† 'Roy. Soc. Proc.,' A, vol. 87 (1912).

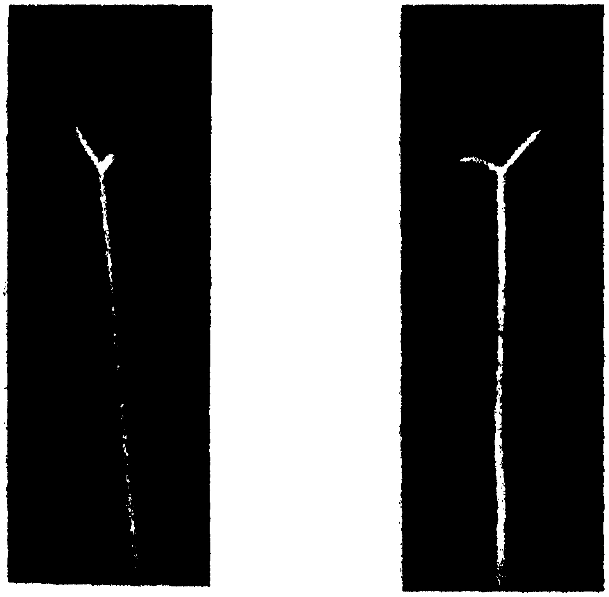


22·5° with the vertical. Both the position and inclination of the outer two should be adjustable, and the distance from the axis is so chosen that the virtual image of an object placed in the line of the axis reflected twice by each pair of mirrors (A, A′, and A″ represent the object, the first, and the second virtual images respectively) comes directly below the object. As can be easily seen from the construction, the pair of mirrors on the right of the axis serve to replace the object with a virtual image tilted by 45° to the left and placed farther away from the lens than the object. The other pair on the left do the same, only with the reversed sense of tilting. Thus, what is seen on the screen is a superposition of two mutually perpendicular views of the same object. In the case of the tracks produced in the expansion apparatus, the object is composed of bright lines on a dark ground instead of a continuous bright surface, and the superposition does not impair the definition of the image. It is always possible to choose an imaginary horizontal line in the middle of the chamber parallel to the mirrors as an axis around which the object space is rotated. When a negative is obtained, a line in the position of this axis will appear as a single line and all in other positions and directions as double lines in general, from the separation and other geometrical data of which the true shape of the object can be deduced. It seems also to be possible to construct a uniformly magnified model of the object in three dimensions directly from the negative, using a similar mirror system as a projecting apparatus. In this case the projected image is the intersection of shadows in space. But, so far, not much actual experiment has been done in this connection.

When many tracks are to be photographed at each expansion, it is better to separate the two images of the object space from each other, so that they occupy different areas on the negative side by side, for then the complication due to overlapping can be avoided. In this case the angle of the outer mirrors with the vertical should be slightly different from 22·5°, and the two companion photographs bear approximately the relation of a plan and a side-view to each other. In either case, the final adjustment of the outer mirrors can be done by putting a cube with white surfaces in place of the expansion chamber.

If the conditions are favourable, one ought to be able to get a thousand or more exposures in an hour. But so far, owing to various difficulties, only about 200 feet of film has been exposed, and the writer cannot at present do more than produce a typical sample of branched tracks photographed with the apparatus.

In the accompanying figure, the two photographs are two mutually perpendicular views of the same track of an α-ray emitted by polonium, taken by the second of the two methods of mirror arrangement described above. In this reproduction the end parts of the track were enlarged 13·2 times from the negative, cut out, and brought together to a convenient distance. The image on the negative itself was 0·42 times the size of the object, so that the true magnification of the reproduction is about 5·5. In this particular exposure



Photograph of a branched α-ray track viewed from two positions at right angles to each other. Actual magnification 5·5.

there was only one track recorded, but usually there are several and it is an easy matter to decide which track in the two images correspond. The slight departure from a straight path noticeable in one of the accompanying photographs is in this case more likely due to an optical error caused by a slight local heterogeneity in the thickness of the gelatine film than to a movement of the air, although it is true that a distortion of this magnitude can arise from the latter cause. The marked smooth curvature in one of the branches cannot be attributed to these disturbances, and is typical of numerous α-ray tracks which finish their range in that way.

In the course of these preliminary experiments, about 3,000 α-ray tracks were photographed and eleven tracks were observed branching like the one shown in the plate, the number of spurred tracks being about thirty-five. Although a far greater number of photographs are required to estimate with accuracy the percentage of such branched tracks, it seems probable that it is not very far from one in a few hundreds.

One of the striking features about the branches is that their shapes and sizes are very similar. The frequent occurrence of such marked branching is very surprising.

With the spurred tracks, the abrupt bending takes place at different distances from the source and the deflection after collision seems to have a wide range, while with the branched tracks the branching takes place near the end of the path and the angle subtended by the branches does not seem to differ much, at any rate, in the small number of cases which have been observed.

No such high percentage of production of swiftly-moving atoms has been observed by the scintillation method, either for recoil or disruption of atoms by collision. This is not contradictory because the expansion method detects moving atoms within the range of the α-rays, which the other method fails to do. At any rate the percentage alone cannot be any clue to the nature of the branching. On the other hand, although a calculation from the geometrical data based upon ordinary laws of dynamics and of the range of particles indicates that the atoms after branching possess nearly equal masses, none of these laws can be relied upon with certainty if disintegration of an atom occurs. It is difficult to speculate with confidence at present about the nature of these rays, and the conclusion must be suspended until sufficient data accumulate.

In conclusion the writer wishes to express his sincere gratitude to Prof. Sir E. Rutherford, who suggested the problem, for his kind encouragement and guidance.

On the Spectra of the Alkaline Earth Fluorides and their Relation to each other.

By SNEHAMOY DATTA, M.Sc. (Calcutta), Research Student, Imperial College.

(Communicated by Prof. A. Fowler, F.R.S. Received April 21, 1921.)

[PLATE 6.]

1. *Introductory.*

It is well known that the limits of the series and the separation of doublets and triplets, which occur in the line spectra of chemically related elements, are intimately connected with the atomic weights of the elements involved. It seemed reasonable to suppose that numerical relations might also exist in the case of band spectra, and the present paper gives an account of an investigation undertaken to determine possible relations among the band spectra of similar compounds of different elements. Previous work by other observers suggested the fluorides of the alkaline earth metals as a suitable group of compounds for the purpose in view. In the first instance a more complete survey of the spectra of these compounds was made, and several bands in addition to those previously described have been observed. These new bands have proved of considerable service in the identification of corresponding bands in the different spectra, and details relating to them are accordingly included in the paper.

The beautiful bands of CaF$_2$, SrF$_2$, and BaF$_2$ were first observed in the flame spectra by Mitscherlich[*] as far back as 1862, but no systematic measurements were made until 1905, when observations of the bands of these compounds in the visible region were made by Fabry[†] with a spectroscope of not very large dispersion. Measurements of the same spectra from photographic plates were first made by Rosch[‡] and further measures were subsequently made by Leopold[§] in the case of SrF$_2$, and by George[||] in the case of BaF$_2$.

The ultra-violet bands of these compounds were studied by Olmsted[¶] in 1906, but the bands of the fluorides were not completely separated from those of the oxides, which in this region are closely grouped with those of the fluorides, particularly in calcium and strontium.

---

[*] A. Mitscherlich, 'Pogg. Ann.,' vol. 121, pp. 459–488 (1864).

[†] Ch. Fabry, 'Astrophys. Journal,' vol. 21, p. 356 (1905).

[‡] J. Rosch, 'Z. Wiss. Photgr.,' vol. 4, p. 384 (1906).

[§] P. Leopold, 'Z. Wiss. Photgr.,' vol. 11, pp. 105, 137 (1912).

[||] H. George, 'Z. Wiss. Photgr.,' vol. 12, p. 237 (1913).

[¶] C. M. Olmsted, 'Z. Wiss. Photgr.,' vol. 4, p. 354 (1906).

The first attempt to discover a relation between the spectra of nearly related compounds was also made by Mitscherlich, who stated in a most general way that the greater the atomic weight the more towards the red the spectrum lies. Later, in 1874, Lecoq* suggested an equation of the type $x = n(1+q)^2$ connecting the bands of $BaCl_2$, $BaBr_2$, and BaI_2, where x = increase of λ between the spectra of the second and third compounds, n = increase of λ between the homologous spectra of the first and second compounds, and $q = B - A/A$, where A = increase of molecular weight from first to second compound, and B that from second to third. In 1894 Hartley,† experimenting on the banded flame spectra of metals, remarked that with an increase in atomic weight the heads of the bands come closer together. In the case of $CaCl_2$, $CaBr_2$, CaI_2, and $SrCl_2$, $SrBr_2$, SrI_2, Olmsted suggested the formula $\sqrt{(\Delta\lambda)}/$molecular weight = constant, where $\Delta\lambda$ is the mean of the differences in the wave-lengths of the first two or three consecutive heads of homologous series in the respective spectra.

Neither the formula of Lecoq nor that of Olmsted is applicable to the spectra of the fluorides under consideration. The new observations, however, have suggested a formula which appears to represent the facts very closely.

2. *Experimental Procedure.*

Observations were chiefly made with the concave grating in the Spectroscopic Laboratory of the Imperial College. The grating has 15,000 lines to an inch over 4 in. ruled surface and is of radius 10 ft. The dispersion obtained on the photographic plate is about 5·5 A. per millimetre in the first order. The negatives for measurements were mostly prepared in the second order, while those for reproduction were taken in the first order and subsequently enlarged.

The band spectra were conveniently produced in the electric arc between carbon poles fed with the various compounds and maintained by an electric main of 110 volts. The arc was extended to a flame by suitable admixture of sodium carbonate with the compounds used, so that the carbon poles could be drawn apart, giving a "flame" having a low order of temperature such as is required for the development of the bands. By carefully moving the arc to and fro, so as to keep only the coloured edge of its image upon the slit, the bands were seen practically free from the line spectrum of the metals, though numerous lines appeared at the centre of the arc. The admixture of sodium carbonate does not inconveniently complicate the spectrum and is a far more convenient method of maintaining a steady flame arc than that of broadening the arc into a wing shape with the help of a magnet, as was adopted by Eagle[*] in his experiments on hydrides.

---

* Lecoq de Boisbaudran, 'Spectres Lumineux,' Paris, 1874.

† W. N. Hartley, 'Trans. Roy. Dublin Soc.,' vol. 7, p. 339 (1901).

Where the bands of the fluorides were free from admixture with those of oxides, they were photographed with the usual iron comparison, and their wave-lengths measured with reference to the iron lines, using Burns's wave-lengths as standards.

Where fluoride and oxide bands were superposed, the fluoride bands were sorted out in the following manner. Three comparison photographs were taken: 1. The fluoride and the oxide. 2. The fluoride and the nitrate. 3. The oxide and the nitrate. From the third plate the common bands and diffuse lines were marked out. These were evidently the bands of the oxides, or the faint diffuse lines of the metal itself. The third plate was taken to make sure of the elimination of the oxide bands, for often it was found that, when carbon poles were used, the oxides gave the more sensitive heads of the fluoride spectra in addition to their own spectra. From each of the first and second plates the common heads were left out, and the others were measured with reference to sharp metallic lines in the comparison oxide spectrum. The common heads omitted are believed to be those due to the oxides, for they were all identified with the common bands of plate 3. The wave-length of the sharp metallic lines—which were taken as reference lines in the measurement of fluoride bands from plates 1 and 2—were then accurately determined from another photograph of the oxides with the iron comparison. Several such plates were taken, and the results of the independent measurements of the wave-lengths of the fluoride bands were in close agreement. The results published here show the mean.

Before proceeding to describe the spectrum of any particular compound, it is best to define the terminology accepted in this paper; and this is more necessary, in view of the fact that, in certain instances, the terminology of band spectra not being completely fixed, misunderstandings arise between different observers.

A *band* is composed in general of a large number of lines, regularly distributed in the spectrum, starting from a line which is the most brilliant of the set, and which is called the *head* of the band. Ordinarily, a band does not occur alone; there are several, of similar character, which usually encroach upon one another, the whole forming a *series* of bands, which can be numerically expressed by certain well-known equations. The manner in which the successive heads occur in a series defines its *structure*. Generally, the different series of similar structure overlap with one another, forming a *group*, but sometimes they are well separated from one another; the different series then form different groups. The series is said to fall off toward the red or toward the violet, according to the direction in which the heads composing the series extend outward from the first head.

---

[*] Eagle, 'Astrophys. Journal,' vol. 30, p. 231 (1909).

Proceeding toward the red, the various series have been designated by the letters A, B, C, with numerical suffixes and dashes. Suffixes have been used to designate similar series standing isolated and dashes when they overlap. Thus A_1, A_1' define two overlapping series of similar structure, forming the group A_1. A_2, A_2' would then stand for a separate group, A_2, of two overlapping series, similar in structure with the series in group A_1, and so on. All these series having similar structure when spoken of generally are designated as those of *class* A. Series of the B class would then refer to another set of similar series, having their structure different from those of class A, and similarly for the series of the C class. When three or four similar series form an overlapping group, they have been designated by successively increasing the number of dashes. Thus B_1, B_2, B_2', B_2'', B_2''' refer to five similar series in CaF_2, one of which, B_1, is isolated from the rest, which form the group B_2 (see Plate 6, fig. III).

In the following Tables, besides the new series of bands, the first two heads of the already known series have been included, as it will be necessary to refer to them in the selection of the homologous series, on the basis of which certain relations have been found.

All the bands—old and new—have been put into the series form—using an equation of the type

$$\nu = a - bm - cm^2, \tag{1}$$

which is really a modification of Deslandres' formula

$$\nu = a - (bm + c)^2, \tag{2}$$

used by the previous observers. The constants of equation (1) are related to those of (2) as follows:—

$$a \text{ of } (1) = a - c^2 \text{ of } (2),$$

$$b \text{ of } (1) = 2bc \text{ of } (2),$$

$$c \text{ of } (1) = b^2 \text{ of } (2).$$

The equation (1) has the advantage that all the constants have a definite meaning; thus a is the end constant, and $= \nu_0$ (the wave number of the first head), $c =$ half the second difference of the wave numbers, and this being a small quantity, b gives approximately the average differences of the wave numbers between the consecutive heads.

3. *The Bands of Magnesium Fluoride.*

Magnesium fluoride presents three groups of bands, which are all in the ultra-violet between λ 3468 and λ 3686, and all fading toward the more refrangible side. The middle group, B_1, is by far the strongest, and consists of four overlapping series of bands. Owing to similarity in structure, these have been called, B_1, B_1', B_1'', B_1''', after the series of the same designation in the bands of CaF_2 and SrF_2. The other two groups are of about equal intensity, and, as there are no corresponding bands in the spectra of the fluorides of the other elements, the more refrangible group has been called D_1, and the less refrangible one E_1 (see fig. I, Plate 6).

In the D_1 group the band heads occur in pairs, and the two series have been named D_1 and D_1'. This group is somewhat similar in appearance to the A_3 or A_4 group of CaF_2, but the presence of a second component to some of the stronger heads marks out a dissimilarity in structure. The E_1 group is more complex, consisting of alternating double and single heads. Proceeding toward the ultra-violet the interval from a doublet to a singlet is rather smaller than that from the singlet to the next doublet. The heads, therefore, appear to occur in groups of three, which have been designated as E_1 and E_1' for the double head and E_1'' for the single one.

Some of the bands of the B_1 and D_1 groups have been previously observed by Olmsted but he did not attempt any grouping into series. It is, however, surprising that the equally intense group E_1 on the less refrangible side was not observed before.

Table I.—Constants for the Calculation of Series in MgF_2.

Series.	$a = \nu_0$.	*b*.	*c*.
D_1	28543 ·38	−20 ·98	+0 ·11
D_1'	28585 ·88	−20 ·91	+0 ·11
B_1	27861 ·28	−28 ·05	+0 ·015
B_1'	27825 ·47	−28 ·85	+0 ·015
B_1''	27814 ·55	−28 ·2601	+0 ·015
B_1'''	27775 ·61	−27 ·57	—
E_1	27149 ·43	−36 ·475	+0 ·175
E_1'	27145 ·44	−36 ·435	+0 ·1786
E_1''	27123 ·57	−36 ·7101	+0 ·1365

Table II.—MgF_2 Bands.

Series.	m.	Datta λ.I.A. (air).	ν (vac.).	$\nu_0 - \nu_c$.	m.	Datta λ.I.A. (air).	ν (vac.).	$\nu_0 - \nu_c$.	Series.	Second heads to D_1'.
D_1	0	3502·44	28543·87	−0·01	0	3503·36	28535·88	0·00	D_1'	3503·30
	1	3499·88	564·24	−0·01	1	00·81	556·66	−0·02		00·62
	2	97·35	584·90	0·00	2	3498·29	577·22	−0·04		3498·10
	3	94·84	605·35	+0·02	3	95·80	597·57	−0·06		95·56
	4	92·41	625·41	−0·13	4	93·31	617·96	+0·20		93·11
	5	90·08	644·43	−1·10	5	90·88	637·95	+0·27		90·61
	6	87·54	665·37	+0·07	6	88·48	657·56	+0·18		88·33
	7	85·17	684·78	−0·07	7	86·11	677·04	+0·18		86·0
	8	82·81	704·21	+0·03	8	83·78	696·21	+0·09		
	9	80·50	723·25	−0·04	9	81·48	715·25	+0·09		
	10	78·33	741·16	−1·02	10	79·19	734·15	+0·17		
	11	76·32	757·86	−2·99	11	76·96	752·56	−0·02		
					12	74·73	771·01	+0·05		
					13	72·55	789·08	−0·04		
					14	70·48	806·24	−0·82		
					15	68·46	823·01	−2·77		
B_1	0	3588·19	27861·28	0·00	0	3592·80	27825·47	0·00	B_1'	
	1	84·59	889·26	−0·05	1	89·11	854·15	+0·05		
	2	80·98	917·36	+0·04	2	85·42	882·80	+0·09		
	3	77·41	945·21	−0·08	3	81·80	911·18	−0·13		
	4	73·79	973·51	+0·11	4	78·06	940·15	−0·26		
	5	70·28	28001·15	−0·14	5	74·48	969·12	−0·22		
					6	71·98	993·95	−2·69		
B_1''	0	3594·22	27814·55	0·00	0	3599·26	27775·61		B_1'''	
	1	90·58	842·74	−0·05	1	95·69	803·18			
	2	86·94	871·00	−0·01						
	3	83·30	899·30	+0·10						
	4	79·66	927·50	+0·09						
	5	76·09	955·53	+0·05						
	6	72·65	983·07	−0·50						
	7	68·0	28011·20	−0·42						
E_1	0	3682·28	27149·43	0·00	0	3682·82	27145·44	0·00	E_1'	
	1	77·36	185·73	0·00	1	77·92	181·69	−0·09		
	2	72·50	221·70	+0·02	2	73·06	217·60	−0·05		
	3	67·69	257·31	+0·05	3	68·26	253·13	+0·02		
	4	62·96	292·58	+0·08	4	63·52	288·83	+0·07		
	5	58·28	327·48	+0·10	5	58·87	323·15	+0·00		
	6	53·57	362·70	+0·79	6	54·25	357·61	−0·01		
	7	49·04	396·66	+0·59	7	49·69	391·72	+0·06		
	8	very faint			8	45·19	425·50	+0·08		
	9	„			8	40·76	458·90	+0·12		
					10	36·42	491·71	0·00		
					11	32·20	524·62	−0·82		
E_1''	0	3685·79	27123·57	0·00	6	3656·73	27338·99	+0·06	E_1''	
	1	80·80	160·14	+0·02	7	52·07	373·86	+0·07		
	2	75·91	196·44	+0·01	8	47·46	408·52	0·00		
	3	71·04	232·47	−0·02	9	42·93	442·91	−0·24		
	4	66·23	268·23	0·00	10	38·41	477·03	−0·27		
	5	61·46	303·71	+0·04	11	33·94	510·87	−0·75		

4. *The Bands of Calcium Fluoride.*

Calcium fluoride shows seven groups of bands, of which five are in the visible spectrum, already well known, and two in the ultra-violet, previously unrecorded. Proceeding toward the red they have been designated as A_1, A_2, A_3, A_4; B_1, B_2; and C_1 respectively, in accordance with the terminology explained before. The new ultra-violet groups are too weak and complex (owing to the superposition of oxide bands), for satisfactory reproduction; only a sketch has therefore been given below to explain their structure (fig. 1).

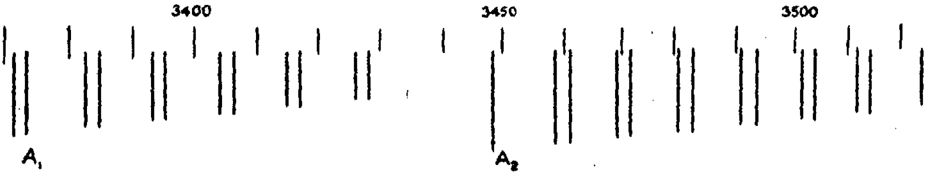


FIG. 1.—Structure of the groups A_1 and A_2 in CaF_2.

The Groups A_1, A_2, A_3 and A_4 are similar in structure as is indicated by their nomenclature. All of them fade toward the red and have double heads (Plate 6, fig. II). Groups B_1 and B_2 fade off toward the violet, the successive intervals between the heads decreasing very rapidly. Group B_1 consists of only one series of weak bands and is remarkable as having the "tail" (marked with black dots) stronger than the head. Group B_2 consists of four overlapping series, the "tails" of two of which (B_2, B_2'') are clearly recognisable (Plate 6, fig. III). Group C_1 consists of two overlapping series C_1, C_1', fading towards the violet. But as the successive intervals between the heads decrease very slowly, their appearance is quite different from those of B.

The B_2 and C_1 groups of bands, though previously known, have been remeasured, as a greater number of heads has been observed in each series. Their series formulæ have also been recalculated.

Table III.—Constants for the Calculation of Series in CaF_2.

Series.	$a = \nu_0$.	b.	c.	Series.	$a = \nu_0$.	b.	c.
A_1	29655·02	+100·61	−1·0	B_1	17094·8	− 6·006	+0·1764
A_1'	29636·74	+101·001	−1·0	B_2	16569·15	− 5·6201	+0·155
A_2	28983·70	+ 90·5002	−0·7	B_2'	16521·72	− 5·4002	+0·155
A_2'	28876·70	+ 90·2001	−0·8	B_2''	16484·87	− 5·6008	+0·155
A_3	19428·5	+ 24·242	+0·1426	B_2'''	16424·27	− 5·2401	+0·155
A_3'	19424·8	+ 24·191	+0·1419	C_1	15978·94	−11·3001	+0·150
A_4	18894·9	+ 20·6018	+0·1029	C_1'	15905·89	−11·0008	+0·120
A_4'	18888·1	+ 20·4766	+0·1017				

Table IV.—CaF_2 Bands.

Series.	m.	Datta λ.I.A. (air).	ν (vac.).	$\nu_0-\nu_c$.	m.	Datta λ.I.A. (air).	ν (vac.).	$\nu_0-\nu_c$.	Series.
A_1	0	3371·15	29655·02	0·00	0	3373·22	29636·74	0·00	A_1'
	1	82·62	555·41	0·00	1	84·64	536·77	+0·03	
	2	93·71	457·85	+0·05	2	95·93	438·61	−0·13	
	3	3404·76	363·28	+0·09	3	3407·02	342·72	−0·02	
	4	15·66	268·62	−0·06	4	18·00	248·57	−0·17	
	5	26·42	176·64	−0·33	5	28·73	156·98	+0·24	
A_2	0	3449·24	28983·66	−0·04	0	3462·02	28876·70	0·00	A_2'
	1	59·95	893·88	−0·02	1	72·78	787·25	−0·05	
	2	70·59	805·33	−0·17	2	83·41	699·35	−0·15	
	3	81·19	718·47	−0·08	3	93·96	612·71	−0·59	
	4	91·60	631·96	−0·94	4	3504·23	528·79	+0·09	
	5	3501·80	548·59	−0·12	5	14·46	445·78	+0·08	
	6	11·84	467·00	+1·1	6	24·54	364·45	+0·15	
A_3 (Fabry A_1)	0	5145·668	19428·5		0	5146·648	19424·8		A_3' (Fabry A_2)
	1	52·246	403·7		1	53·229	400·0		
	...	......	......		...	......	......		
	18	5276·668	18946·2		18	5277·583	18942·9		
A_4 (Fabry B_1)	0	5290·993	18894·9		0	5292·899	18888·1		A_4' (Fabry B_2)
	1	96·882	873·9		1	98·623	867·7		
	...	......	......		...	......	......		
	20	5420·994	441·8		20	5422·112	438·0		
B_1 (Fabry C)	0	5848·142	17094·8						
	1	46·258	100·6						
	...	......	......						
	10	33·635	137·3						
Datta B_2 (Fabry D)	0	6037·29	16559·15	0·00	0	6050·97	16521·72	0·00	Datta B_2' (Fabry D′)
	1	35·30	564·60	−0·01	1	49·03	527·01	+0·04	
	2	33·40	569·82	+0·05	2	47·22	531·96	+0·06	
	3	31·62	574·71	+0·09	3	45·58	536·44	−0·08	
	4	29·96	579·21	+0·06	4	43·98	540·82	−0·02	
	5	28·43	583·47	+0·10	5	42·56	544·71	−0·18	
	6	27·07	587·21	−0·08	6	41·23	548·35	−0·23	
	7	25·78	590·76	−0·13	7	39·96	551·99	+0·07	
	8	24·55	594·16	−0·03					
Datta B_2'' (Fabry D″)	0	6064·5	16484·87	0·00	0	6086·87	16424·27	0·00	Datta B_2''' (Fabry D‴)
	1	62·47	490·38	+0·07	1	84·99	429·34	−0·01	
	2	60·57	495·52	+0·07	2	83·22	434·12	−0·01	
	3	58·81	500·34	+0·06	3	81·56	438·58	−0·02	
	4	57·14	504·89	+0·10	4	80·04	442·72	−0·03	
	5	55·62	509·04	+0·04	5	78·66	446·45	−0·15	
	6	54·23	512·82	−0·07	6	77·53	449·50	−0·63	
	7	52·93	516·37	−0·11	7	76·23	453·05	−0·31	
	8	51·93	519·10	−0·65	8	74·81	456·87	+0·60	
					9	73·84	459·52	+0·65	

Table IV.—CaF_2 Bands—(*contd.*).

Series.	*m.*	Datta λ.I.A. (air).	ν (vac.).	$\nu_0 - \nu_c$.	*m.*	Datta λ.I.A. (air).	ν (vac.).	$\nu_0 - \nu_c$.	Series.
Datta C_1 (Fabry E)	0	6256 ·52	15978 ·94	0 ·00	0	6285 ·25	15905 ·89	0 ·00	Datta C_1' (Fabry E′)
	1	52 ·14	990 ·12	+0 ·08	1	80 ·95	916 ·75	−0 ·02	
	2	47 ·94	16000 ·86	−0 ·08	2	76 ·72	927 ·49	+0 ·08	
	3	43 ·72	011 ·68	+0 ·19	3	72 ·67	937 ·78	−0 ·08	
	4	39 ·77	021 ·82	+0 ·08	4	68 ·57	948 ·20	+0 ·23	
	5	36 ·08	031 ·42	−0 ·27	5	64 ·68	958 ·10	+0 ·21	
	6	32 ·28	041 ·07	−0 ·27	6	60 ·96	967 ·58	+0 ·01	
	7	28 ·56	050 ·64	−0 ·05	7	57 ·37	976 ·74	−0 ·27	
	8	25 ·47	058 ·61	−1 ·13	8	54 ·12	985 ·07	−1 ·14	
	9	22 ·54	066 ·16	−2 ·83	9	50 ·31	994 ·80	−0 ·87	
					10	46 ·53	16004 ·48	+0 ·59	

5. *The Bands of Strontium Fluoride.*

Strontium fluoride shows eleven groups of bands, of which nine are in the visible spectrum and two in the ultra-violet. Excepting one in the visible region (designated as B_2, Plate 6, fig. IV) and the two in the ultra-violet (designated as A_1 and A_2), all were previously known. Proceeding toward the red, these known groups are designated as A_3, A_4, B_1, B_2, B_3, C_1, C_2, C_3, and C_4 repectively (Plate 6, figs. IV and V). The new ultra-violet groups are too weak for satisfactory reproduction, and also complex owing to the superposition of oxide bands; only a sketch has therefore been given below to explain their structure (fig. 2).

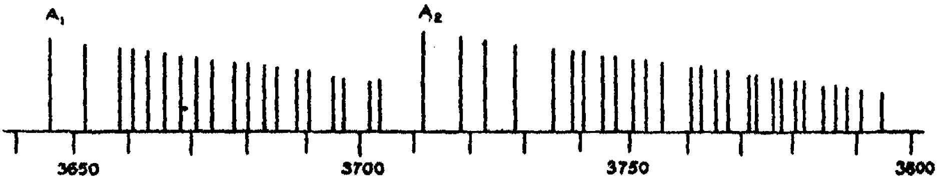


Fig. 2.—Structure of the groups A_1 and A_2 in SrF_2.

The groups A_1, A_2, A_3 and A_4 are similar in structure, all of them fading toward the red. These have double heads and are similar to the corresponding groups in CaF_2. Groups of the B and C class fade toward the violet, but the structure of the B class is quite different from that of the C class. As in the corresponding groups in CaF_2, the successive intervals between the heads decrease very rapidly in the B class, these having well pronounced "tails." The tails of B_1 and B_3 have been marked with dots (Plate 6, fig. IV). The group B_2 is very weak and has been photographed with 30 minutes' exposure.

Table V.—Constants for the Calculation of Series in SrF_2.

Series.	$a = \nu_0$.	b.	c.	Series.	$a = \nu_0$.	b.	c.
A_1	27417 ·16	+ 44 ·35	−0 ·22	B_3	15578 ·921	− 3 ·3203	+ 0 ·0443
A_1'	27309 ·72	+ 44 ·70	−0 ·12	C_1	15356 ·300	− 7 ·8989	+ 0 ·0592
A_2	26929 ·35	+ 41 ·001	−0 ·20	C_1'	15319 ·617	− 8 ·0935	+ 0 ·0592
A_2'	26740 ·04	+ 40 ·801	−0 ·41	C_2	15076 ·973	− 7 ·8733	+ 0 ·0549
A_3	17764 ·15	+ 11 ·732	+ 0 ·0417	C_2'	15025 ·049	− 7 ·6006	+ 0 ·0542
A_3'	17764 ·15	+ 12 ·0204	+ 0 ·0391	C_3	14860 ·720	−12 ·4036	+ 0 ·07472
A_4	17325 ·104	+ 9 ·4831	+ 0 ·0347	C_3'	14835 ·25	−12 ·085	+ 0 ·05166
A_4'	17302 ·615	+ 8 ·9796	+ 0 ·04405	C_4	14557 ·99	−11 ·769	+ 0 ·03753
B_1	15857 ·83	− 3 ·4908	+ 0 ·0531	C_4'	14557 ·99	−11 ·9023	+ 0 ·05419
B_2	15787 ·93	− 3 ·7001	+ 0 ·0901				

Table VI.—SrF_2 Bands.

Series.	m.	λ.I.A. (air).	ν (vac.).	$\nu_0 - \nu_c$.	m.	λ.I.A. (air.)	ν (vac.).	$\nu_0 - \nu_c$.	Series.
Datta A_1	0	3646 ·29	27417 ·16	0 ·00	0	3660 ·66	27309 ·72	0 ·00	Datta A_1'
	1	52 ·17	373 ·03	0 ·00	1	66 ·64	265 ·20	+ 0 ·06	
	2	58 ·01	329 ·27	−0 ·07	2	72 ·63	220 ·73	−0 ·07	
	3	63 ·82	285 ·95	−0 ·14	3	78 ·59	176 ·65	−0 ·05	
	4	69 ·59	243 ·06	−0 ·22	4	84 ·54	132 ·70	−0 ·14	
	5	75 ·27	200 ·96	0 ·00	5	90 ·42	089 ·48	+ 0 ·26	
	6	80 ·97	158 ·85	−0 ·13	6	96 ·26	046 ·69	+ 0 ·85	
	7	86 ·56	117 ·62	+ 0 ·13	7	3702 ·15	003 ·67	+ 0 ·97	
	8	92 ·13	076 ·79	+ 0 ·20					
	9	97 ·70	035 ·95	+ 0 ·12					
	10	3703 ·12	26996 ·38	+ 0 ·72					
Datta A_2	0	3712 ·36	26929 ·35	0 ·00	0	3738 ·65	26740 ·04	0 ·00	Datta A_2'
	1	18 ·02	888 ·45	−0 ·10	1	44 ·31	699 ·63	−0 ·02.	
	2	23 ·60	848 ·09	−0 ·06	2	49 ·88	659 ·99	−0 ·09	
	3	29 ·16	808 ·14	−0 ·01	3	55 ·58	619 ·46	−1 ·87	
	4	34 ·66	768 ·60	+ 0 ·05	4	60 ·81	582 ·52	−0 ·91	
	5	40 ·02	730 ·32	+ 0 ·97	5	65 ·99	545 ·90	+ 0 ·29	
	6	45 ·56	690 ·66	+ 0 ·11	6	71 ·09	510 ·08	+ 0 ·08	
	7	50 ·98	652 ·17	+ 0 ·02	7	76 ·14	474 ·57	+ 0 ·04	
	8				8	81 ·07	440 ·06	+ 0 ·18	
	9	61 ·66	576 ·45	−0 ·10	9	85 ·79	406 ·34	+ 1 ·08	
	10	66 ·93	539 ·28	−0 ·07	10	90 ·710	372 ·84	−0 ·22	
	11	72 ·0	503 ·62	+ 1 ·07	11	95 ·55	339 ·22	−1 ·68	
	12	77 ·29	466 ·51	+ 0 ·36					
	13	82 ·44	430 ·48	+ 0 ·88					
	14	87 ·75	393 ·37	−1 ·16					
	15	92 ·93	357 ·41	−1 ·94					
A_3 (Leopold A_1)			(air)				(air)		A_3' (Leopold A_2)
	0	5629 ·316	17764 ·148		0	5629 ·316	17764 ·148		
	1	33 ·086	752 ·259		1	33 ·086	752 ·259		
	…	……	……		…	……	……		
	29	5750 ·57	389 ·58		29	5752 ·57	383 ·53		

Table VI.—SrF_2 Bands—(*contd.*).

Series.	m.	λ.I.A. (air).	ν (vac.).	$\nu_0 - \nu_c$.	m.	λ.I.A. (air).	ν (vac.).	$\nu_0 - \nu_c$.	Series.
A_4 (Leopold B_1)	0	5771·971	17825·104		0	5779·473	17802·615		A_4' (Leopold B_2)
	1	74·822	816·551		1	82·169	294·548		
	...	......	......		...	......	......		
	9	98·511	245·806		26	5859·542	066·180		
							(vac.)		Datta B_2.
B_1 (Leopold D)	0	6306·032	15857·832		0	6332·20	15787·98	0·00	
	1	04·721	861·130		1	30·74	791·57	+0·08	
	...	......	......		...	......	......		
	23	6285·02	910·85		2	29·39	794·94	−0·08	
					3	28·11	798·14	−0·08	
					4	26·88	801·21	−0·08	
B_3 (Leopold D″)	0	6418·927	15578·921		5	25·72	804·10	−0·08	
	1	17·687	581·937		6	24·60	806·90	+0·01	
	...	......	......		7	23·56	809·50	+0·08	
	24	6397·406	631·335		8	22·54	812·08	+0·31	
					9	21·55	814·53	+0·59	
							(air)		
C_1 (Leopold E)	0	6511·985	15356·300		0	6527·578	15319·617		C_1' (Leopold E′)
	1	08·676	364·107		1	24·158	327·648		
	...	......	......		...	......	......		
	15	6467·70	461·45		7	04·71	373·48		
C_2 (Leopold E″)	0	6632·631	15076·978		0	6655·552	15025·049		C_2' (Leopold E‴)
	1	29·324	084·494		1	52·091	032·867		
	...	......	......		...	......	......		
	15	6588·199	178·655		6	35·562	070·313		
C_3 (Leopold F)	0	6729·15	14860·72		0	6740·70	14835·25		C_3' (Leopold F′)
	1	23·80	872·54		1	35·16	847·46		
	...	......	......		...	......	......		
	8	6687·13	954·10		11	6688·18	962·94		
C_4 (Leopold F″)	0	6869·08	14557·99		0	6869·08	14557·99		C_4' (Leopold F‴)
	1	63·39	570·06		1	63·39	570·06		
	...	......	......		...	......	......		
	19	6773·33	763·79		19	6772·98	764·66		

6. *The Bands of Barium Fluoride.*

Barium fluoride shows four groups of bands, which are all in the visible region between λ 4937 and λ 5160 and were already well known. Proceeding toward the red, they were designated as A, B, C, and E by George, but their nomenclatures have been altered to B_1, B_2, C_1, and B_3 respectively, in accordance with the groups of CaF_2 and SrF_2, with which they are believed to be homologous.

Although the bands of BaF_2 fade in the opposite direction, *i.e.*, towards the red, the structure of the B_1, B_2, C_1, and B_3 groups are similar to the corresponding groups in CaF_2. Their "tails" are likewise clearly recognisable and have been marked with dots (Plate 6, fig. VI).

The group designated as D by George was never observed. In this connection it may be mentioned that unless the arc is very well supplied with the salt (BaF_2) fine flutings always appear as a sort of background to the stronger bands of BaF_2. These become very weak when the arc is well supplied with BaF_2 and they strengthen with BaO. It is therefore quite likely that these bands are due to the oxides.

In the ultra-violet there appears a sharp head at λ 3810, followed by a series of fine flutings, which have no homologues elsewhere. It has been previously observed by Olmsted.

Table VII.—Constants for the Calculation of Series in BaF_2.

Series.	$a = \nu_0$.	*b.*	*c.*
B_1	20253·837	+ 6·1092	−0·1838
B_2	20198·75	+ 6·5375	−0·1695
C_1	20031·41	+ 11·5539	−0·14908
C_1'	19997·35	+ 11·9478	−0·1278
B_3	19534·47	+ 8·3066	−0·15404

Table VIII.—BaF_2 Bands.

Series.	*m.*	λ (air).	ν (air).	*m.*	λ (air).	ν (air).	Series.
B_1 (George A)	0	4937·336	20253·837	0	4950·802	20898·75	B_2 (George B)
	1	38·812	47·780	1	52·612	91·37	
	...	......	......	...	......	......	
	9	47·315	12·98	18	71·70	18·84	
C_1 (George C)	0	4992·146	20031·41	0	5000·662	19997·35	C_1' (George C′)
	...	......	......	...	......	......	
	8	5012·859	19948·70	29	52·307	753·84	
B_3 (George E)	0	5119·157	19534·47				
	1	21·296	526·31				
	...	......	......				
	17	43·949	440·32				

7. *Remarks on the CaF_2 Bands. Orange Group B_2 (Series B_2, B_2', B_2'', B_2''').*

The series of bands designated H_1, H_2, H_3, H_4, and assigned to SrF_2 by Leopold, and those called F_1, F_2, F_3 by George, and attributed by him to BaF_2, have not been included in the tables of Sr and Ba, for the simple reason that they are really due to CaF_2. A comparison of the measurements of Leopold and George with each other and with the wave-lengths of the orange series of CaF_2 given by Rosch, as in the following Table, clearly show their identity:—

m.	Rosch—CaF_2. λ I.A. (air).	Leopold—SrF_2. λ I.A. (air).	George—BaF_2. λ I.A. (air).
	D (B_2).	H_1.	F_1.
0	6037·153	6036·908	6036·869
1	35·039	34·831	34·820
2	33·251	32·934	32·875
3	31·327	31·099	31·083
4	29·654	29·451	29·400
	D′ (B_2').	H_2.	F_2.
0	6051·084	6050·810	6050·705
1	48·947	48·78	48·706
2	47·156	46·97	46·860
3	45·474	45·10	45·055
4	43·939	43·85	43·439
	D″ (B_2'').	H_3.	F_3.
0	6064·612	6064·395	6064·380
1	62·627	62·344	62·310
2	60·643	60·431	60·378
3	58·880	58·653	58·581
4	57·192	57·026	56·934
5	55·668	55·457	55·412
6	54·295	54·14	54·010
	D‴ (B_2''').	H_4.	
0	6084·985	6084·24	
1	81·565	81·51	
2	80·040	79·53	
3	76·225	76·98	
4	74·810	75·24	

All the different observers agree that with the increase of molecular weight the homologous bands are shifted towards the red. Consequently it seems improbable that exactly similar groups of bands of CaF_2, SrF_2, and BaF_2 would occupy the same position. If SrF_2 and BaF_2 have really similar groups, we should expect them in the red and infra-red respectively.

With these doubts, the following observations were made, which clearly show that the group in question really belongs to CaF_2.

Several comparison photographs were taken in which CaF_2 formed the outer spectrum and either SrF_2 or BaF_2 formed the inner one. This group was then found to be identical in position with the orange group B_2 of CaF_2, but of much lower intensity.

This orange group is the most sensitive in the spectrum of CaF_2. The arc between hard carbon rods, which contain very little CaF_2, shows this group faintly, and it becomes more intense when the arc is fed by a compound of calcium other than the fluoride. Thus a few plates of CaO spectrum were obtained where this group was superposed on the spectrum of the oxide. It is very likely that the appearance of this group in the spectrum of SrF_2 and BaF_2 is due to CaF_2 present in the carbon rod itself. BaF_2 in a copper arc or in the oxy-coal gas flame did not show it at all. Hence it is believed that this group really belongs to CaF_2.

8. *The Relation between the Spectrum of the Alkaline Earth Fluorides.*

Possible relations among the band spectra of nearly related compounds are to be expected from the occurrence of similar series in the different spectra. In the spectra of MgF_2, CaF_2, and SrF_2, such series can be picked out quite readily. The structure of the series and its position relative to the molecular weight of the compound, assisted by the direction in which it shades off, give us the best guidance for such selection. The following three classes have thus been identified:—

Class A—

(*a*) The ultra-violet groups A_1, A_2 of CaF_2.
(*b*) The blue-green and the green groups A_3, A_4 of CaF_2.
(*c*) The ultra-violet groups A_1, A_2 of SrF_2.
(*d*) The green and the yellow groups A_3, A_4 of SrF_2.

These have the following similarity:—All of them fade toward the red; the bands of SrF_2 are more toward the red than the corresponding bands of CaF_2; they present the appearance of pair series, a head of A_1 with one of A_1', a head of A_2 with one of A_2', and so on (figs. II and IV).

Class C—

(*a*) The red group C_1 of CaF_2.
(*b*) The red groups C_1, C_2, C_3, and C_4 of SrF_2.

All these shade off towards the violet. The bands of SrF_2 are further in the red than those of CaF_2; they present the same fluted appearance, and the arrangement is also similar, for in each group we find the second series to be in advance of the first by a few heads, C_1' in advance of C_1, C_2' of C_2, etc.

Class B—

(*a*) The strong ultra-violet group B_1 of MgF_2.
(*b*) The yellow-green group B_1 of CaF_2.
(*c*) The orange group B_2 of CaF_2.
(*d*) The red groups B_1, B_2, and B_3 of SrF_2.

These have the following similarity:—They all fade towards the red and form the most intense group in the respective spectra (the group B_1 of CaF_2 and the group B_2 of SrF_2 are, however, weak); their positions are consistent with their increasing molecular weight and they all have a typical fluted structure totally different from those of class A and C.

In this classification, BaF_2 bands present an anomaly as to their positions in the blue region, but on the similarity of structure the groups B_1, B_2, B_3 and C_1 respectively appear to be homologous with the B_1, B_2, B_3 and C_1 groups of SrF_2.

The classification of practically all the series in the spectra of the alkaline earth fluorides in the above way provides a basis for the further investigation of possible relationship between the characteristics of the various series of bands and the molecular weights of the respective compounds. Of the numerous formulæ tried, the one of the two which represents the facts of observation most closely, is

$$\frac{\log b \times \sqrt{(m')}}{a} = \text{const.},$$

where a and b are the numerical values of the constants used in calculating the series and m' is the molecular weight of the compound.

The 4th column of the Table on opposite page shows how closely the formula applies to the different groups in each class.

We thus find that in each class the values of $\frac{\log b \times \sqrt{(m')}}{a}$ agree well with each other, although varying from one class to another. It is a remarkable fact that the value for class C is very close to that for class A. This is what we should expect when we notice that the series of the C class resemble those of the A class, both fading off very slowly, though in opposite directions.

In seeking relations between the nearly related compounds, some function of the molecular number—a conception of Dr. H. S. Allen* which means the sum total of the nucleus positive charge in the molecule and which plays a part like the atomic number in the elements—was thought worth trying, and hence a separate column introducing molecular number M instead of molecular weight m', has been calculated. This gives a better agreement between the values as given in the 5th column of the preceding Table when instead of $\sqrt{(M)}$, log M is used. The logarithmic function of the atomic number has best answered the relations between the allied elements; hence the introduction of log M seems to be on the right path, and the formula adopted becomes

$$\frac{\log b \times \log M}{a} = \text{constant}.$$

---

* Dr. H. S. Allen, 'Phil. Mag.,' vol. 35 (1918).

Homologous series.	$a = \nu_0$.	b.	$\frac{\log b \times \sqrt{(m')}}{a}$.	$\frac{\log b \times \log M}{a}$.
CaF_2 A_1	29655·02	100·61	$5·964 \times 10^{-4}$	$10·67 \times 10^{-5}$
" A_1'	29636·74	101·001	5·973	10·68
" A_2	28983·70	90·5002	5·962	10·66
" A_2'	28876·70	90·2001	5·979	10·69
" A_3	19428·5	24·242	6·298	11·26
" A_3'	19424·8	24·191	6·290	11·25
" A_4	18894·9	20·6018	6·141	10·98
" A_4'	18888·1	20·4766	6·130	10·99
SrF_2 A_1	27417·16	44·35	6·731	10·51
" A_1'	27309·72	44·70	6·772	10·56
" A_2	26929·35	41·001	6·711	10·47
" A_2'	26740·04	40·801	6·750	10·53
" A_3	17764·15	11·782	6·746	10·63
" A_3'	17764·15	12·0204	6·812	10·63
" A_4	17325·1	9·4831	6·206	9·86
" A_4'	17302·61	8·9796	6·012	9·63
CaF_2 C_1	15978·94	11·3001	$5·782 \times 10^{-4}$	$10·34 \times 10^{-5}$
" C_1'	15905·89	11·0003	5·820	10·41
SrF_2 C_1	15356·80	7·8989	6·550	10·21
" C_1'	15319·617	8·0935	6·572	10·36
" C_2	15076·978	7·8733	6·660	10·39
" C_2'	15025·049	7·6009	6·569	10·36
BaF_2 C_1	20031·41	11·9478	7·126	10·05
" C_1'	19997·35	11·5539	7·048	9·93
MgF_2 B_1	27861·28	28·05	$4·103 \times 10^{-4}$	$7·67 \times 10^{-5}$
" B_1'	27825·47	28·85	4·143	7·75
" B_1''	27814·55	28·2601	4·120	7·70
" B_1'''	27775·61	27·57	4·012	7·12
CaF_2 B_1	17094·8	6·006	4·022	7·19
" B_2	16559·15	5·6201	3·998	7·15
" B_2'	16521·72	5·4002	3·915	7·00
" B_2''	16484·87	5·6003	4·007	7·17
" B_2'''	16424·27	5·2401	3·867	6·92
SrF_2 B_1	15857·83	3·4908	3·836	5·99
" B_2	15787·98	3·7001	4·033	6·29
" B_3	15578·921	3·3208	3·721	5·85
BaF_2 B_1	20258·837	6·5375	5·335	7·52
" B_2	20198·75	6·1092	5·109	7·20

The molecular numbers of the compounds used are:—

$MgF_2 = 12 + 2 \times 9 = 30$, $SrF_2 = 38 + 2 \times 9 = 56$,
$CaF_2 = 20 + 2 \times 9 = 38$, $BaF_2 = 56 + 2 \times 9 = 74$.

These formulæ thus give relations between the wave number of the first head in the series, the average difference in wave number between the consecutive heads, and the molecular weight or molecular number. The fact that the calculated values of the above relations come out as constants may be interpreted in the following way:—The mechanism of emission of bands of the similar compounds of the same periodic group is essentially the same, they emit the same types of bands, so that if the system owing to their possessing greater energy or lesser mass appear on the high frequency side, the successive band heads will be more widely separated, *i.e.*, will have a greater frequency difference, and if owing to either lesser energy or greater mass they appear on the low frequency side, the successive heads will be closer together.

Beryllium fluoride bands, previously unknown, have also been observed, and, from their positions and separations, it is expected that they will agree with the above formula. It is hoped to discuss them in a separate paper.

9. *A New Relation.*

A striking feature of the series of the B class is that they show a strong line on the side in which they fade off, and this may be mistaken for the head of the series.* It is, however, possible to arrive at an explanation of the appearance of this line in the following way. In all these series the successive intervals of the heads gradually decrease, till finally the heads come too close to be resolved (see figs IV, etc.). The strong line on the shading side, which has been named as the "tail" of the series, may thus be regarded as due to the crowding of several unresolved heads in a finite space. The intensity of the tail would therefore depend on the number of heads crowding, and their respective intensity. The falling off of intensity in the B_1 series of CaF_2 is less gradual than that in its B_2 series, and, at the same time, the rate of decrease of intervals between the successive heads is faster in the former than in the latter. Both these factors thus contribute to the crowding of stronger lines to form the tail of the B_1 series. Hence, although the general intensity of the B_2 series is greater, the tail of B_1 appears much more striking than that of B_2, and gives the false appearance of the head.

The B class series, with diminishing intervals, are all expressed by the equation $\nu = a - bm - cm^2$, in which b and c have values of opposite sign. Thus, for the nearest positive integral value of m, satisfying the condition $\partial\nu/\partial m = -b - 2cm = 0$, ν is a maximum or minimum depending on the sign of c. Near this value of m, since $\partial\nu/\partial m = 0$, there will be a crowding of lines. Theoretically, therefore, from a knowledge of b and c, the position of the tail, *i.e.*, the turning point, can be predicted, and the following Tables will show the closeness of the predicted values to the observed ones :—

---

* With regard to the B class of bands in SrF_2, it may be remarked that two different interpretations have been given (Konen, 'Das Leuchten der Gase und Dämpfe,' pp. 238–241). Fabry, whose interpretations have been accepted in the present paper, has described the first line towards the red as the head of the series. Deslandres, on the other hand, took the strong line towards the violet—the one regarded as tail here—as the head. Konen suggested that as the strong line (tail) may possibly be due to a crowding of several heads, Fabry's interpretation is, perhaps, the more correct one.

Series.	m for which $d\nu/dm = 0 = -b/2c$.	$\nu_0 \sim \nu$ tail (cal.).	$\nu_0 \sim \nu$ tail (obs.).	Remarks.
CaF_2 B_1	17	51·12	49·6	The difference between the observed and the calculated value is within the limits of variations of the values of b and c so as to approximately satisfy the equation $\nu = a - bm - cm^2$.
" B_2	18	50·94	48·7	
" B_2''	18	50·58	49·4	
SrF_2 B_1	33	57·37	58·5	
" B_3	36	60·8	57·4	
BaF_2 B_3	27	111·7	111·0	

For values of m equally greater or less than this critical value, the values of ν are almost equal, so that, with a low dispersion, the two heads practically coincide. With a high dispersion, however, a fainter head, corresponding to the greater value of m, is to be expected by the side of the stronger head, corresponding to the lower value of m.[*] For $m = (-b/c)$, $\nu_m = \nu_0$, and, for values exceeding $(-b/c)$, ν_m will be less than ν_0; if, therefore, the successive heads fall in intensity so slowly that the bands $m = -b/c+1$, $-b/c+2$, etc., are not very weak, we should expect one or more fainter heads, whose frequencies are less than that of the head ν_0. In the B_1 series of CaF_2, since the intensities fall off rather slowly, the thirty-fifth or the thirty-sixth band, corresponding to $m = -b/c+1$, $-b/c+2$ respectively, probably survive in intensity, and appear on the low frequency side of the head (ν_0). This, perhaps, explains the appearance of the two faint bands in front of the strong head in fig. III.

Another interesting fact observed is that the difference in frequency $(\Delta\nu)$ between the head and tail in the homologous compounds is nearly constant for different series of the same compound, but varies from one to another, such that $\log \Delta\nu = K \log M$, where M is the molecular number. Thus :—

Series.	$\Delta\nu$.	M.	Log $\Delta\nu$.	Log M.	K.
CaF_2 B_1	49·6	38	1·695	1·58	1·06
" B_2	48·7	38	1·687	1·58	1·06
" B_2''	49·4	38	1·693	1·58	1·06
SrF_2 B_1	58·5	56	1·767	1·748	1·01
" B_3	57·4	56	1·759	1·748	1·01
BaF_2 B_2	111·68	74	2·048	1·869	1·09
" B_3	111·8	74	2·0484	1·869	1·09

---

[*] Depending on the values of B and C, they may exactly coincide, in which case their separation is impossible.

The C class of series also should have tails, but, from the theoretical calculation, it is found that, in the C_1' series of CaF_2, the thirty-eighth band would correspond to the tail; as the series is very weak, the heads about the thirty-eighth band are too weak to give the impression of a tail. Though the corresponding series of SrF_2 is very strong, the tail would occur at the sixty-seventh band, and so the heads forming the tail are weak, and do not, therefore, show any clear tail.* But, even in this class, $\Delta\nu$ and M exhibit some indication of the previous relation, thus :—

Series.	m for max.	$\Delta\nu$ (calculated).	M.	Log $\Delta\nu$.	Log M.	K.
CaF_2 C_1'	38	212·6	38	2·327	1·58	1·47
SrF_2 C_1'	67	263·3	56	2·419	1·748	1·39
BaF_2 C_1'	47	290·2	74	2·462	1·869	1·35

The series C_1' of BaF_2 shows a tail, as is to be expected, considering the general intensity of the group. The observed $\Delta\nu = 296{\cdot}8$ closely agrees with the calculated one.

10. *Summary.*

1. Nine series of bands have been found in the spectrum of MgF_2.

2. Four new series have been found for CaF_2 in the ultra-violet, and additional heads recorded in some of the series in the visible region.

3. Four new series have been found for SrF_2 in the ultra-violet and a very weak one in the red.

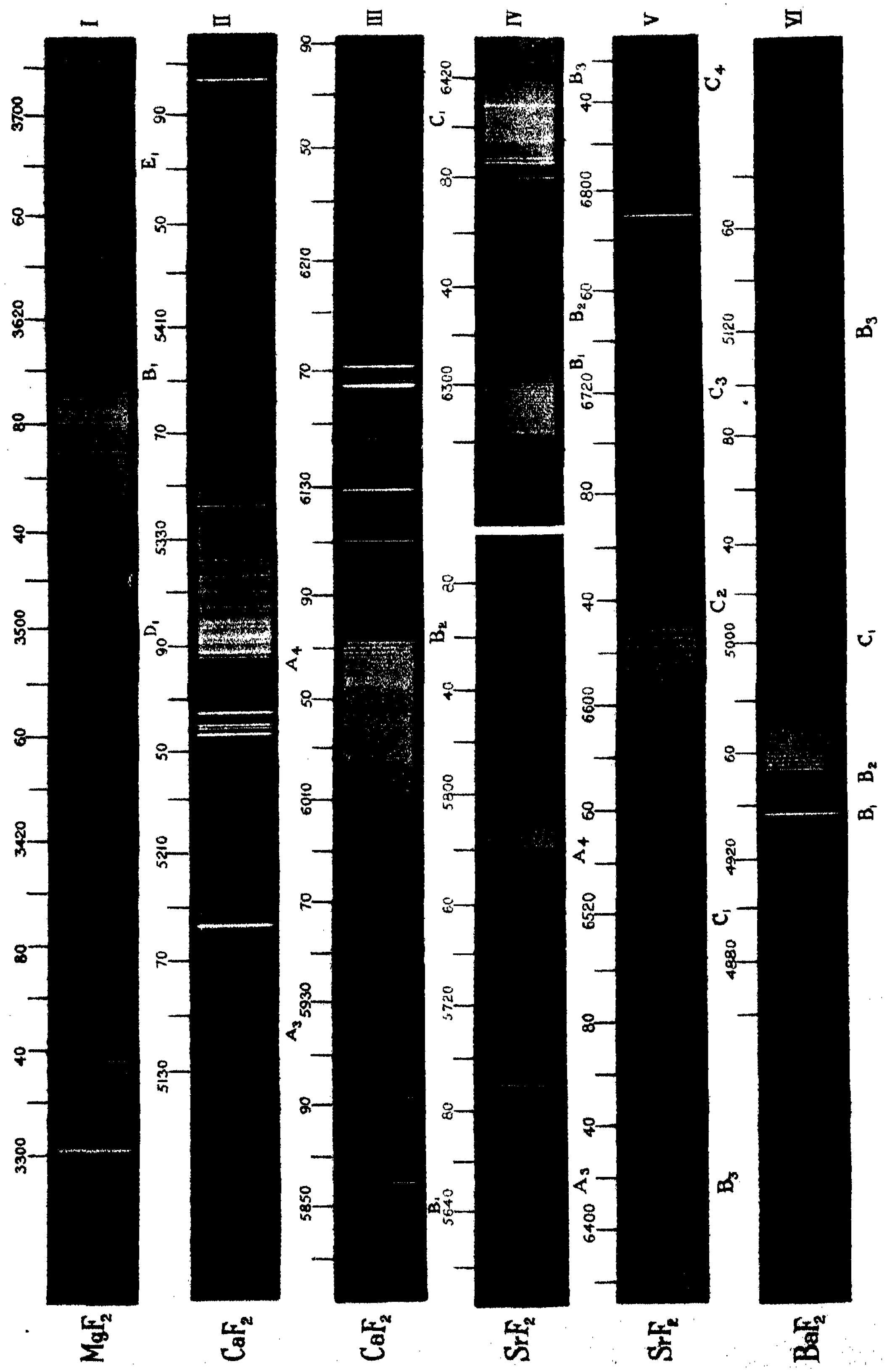
4. Series designated as H_1, H_2, H_3, H_4, and ascribed to SrF_2 by Leopold, also the same series called F_1, F_2, F_3, and allotted to BaF_2 by George, have been proved to be really due to CaF_2, and corresponding with its series now designated as B_2, B_2', B_2'', B_2'''.

5. The homologous series of the different alkaline earth fluorides have been connected by empirical equations, using the constants of the series equations and the molecular weight or the molecular numbers of the respective compounds.

6. An explanation has been given of the appearance of a "tail" in some of the bands, starting with their series equations. It has also been shown that the difference in wave numbers of the heads and tails of the similar series is constant for the same compound, but varies from one to another in a definite way.

---

* In the series C_1 and C_1' of SrF_2, there are indications of faint tails near about the calculated positions; in C_2'' and C_2''' the calculated positions of the tails fall on the preceding series C_1 and C_1' respectively, hence they cannot possibly be detected.


I
II
III
IV
V
VI
MgF_2
CaF_2
CaF_2
SrF_2
SrF_2
BaF_2
3700
60
3620
80
40
3500
60
3420
80
40
3300
90
E₁
50
5410
B₁
70
5330
D₁
90
50
5210
70
5130
90
50
6210
70
6130
90
A₄
50
6010
70
A₃ 5930
90
5850
C₁ 6420
80
40
6300
80
B₂
40
5800
60
5720
80
B₁ 5640
B₃
40
6800
B₂ 60
B₁ 6720
80
40
6600
60
A₄
6520
80
40
A₃
6400
C₄
60
5120
C₃
80
40
C₂
5000
60
4920
C₁
4880
B₃
B₃
C₁
B₂
B₁

In closing, I beg to thank Prof. A. Fowler, F.R.S., for the very valuable suggestions he was always most ready to give, and for the trouble he took in revising the paper, and also to Mr. W. Bradfield, for his help in producing the enlargements.

DESCRIPTION OF PLATE 6.

Fig. I.—Spectrum of MgF_2, region about λ 3460–λ3700. On the left is the group D_1, next comes B_1, and finally on the right there is the group C_1.

Fig. II.—Spectrum of CaF_2, region about λ 5145–λ 5430, having the group A_3 on the left and the group A_4 on the right.

Fig. III.—Spectrum of CaF_2, region about λ 5835–λ 6285. The group B_1 is on the left, the dot at about λ 5833 marks its "tail." The group B_2 comes next. The dot at about λ 6022 marks the "tail" of the series B_2, and that at about λ 6050 shows the "tail" of B_2''. On the right there is the group C_1. The dots are not very clear in the reproduction.

Fig. IV.—Spectrum of SrF_2 λ 5620–λ 6420. Proceeding towards the red, there are the groups A_3, A_4, B_1, B_2, and B_3 in succession. The dot at about λ 6284 marks the "tail" of the series B_1, and that at about λ 6394 shows the "tail" of the series B_3.

Fig. V.—Spectrum of SrF_2 λ 6380–λ 6870, showing the groups B_3, C_1, C_2, C_3, and C_4, from left to right.

Fig. VI.—Spectrum of BaF_2 λ 4937–λ 5160, showing the group B_1, B_2, C_1, and B_3, in succession from left to right. The dots mark the positions of the tails of the respective series.

The Absorption of Light by Elements in a State of Vapour: The Halogens.

By Sir J. J. DOBBIE, F.R.S., and J. J. FOX, O.B.E., D.Sc.

(Received June 4, 1921.)

[PLATE 7.]

In previous papers* we have given an account of the effect produced on the light of the Nernst filament by passing it through the vapour of the elements of the sulphur, phosphorus and mercury groups and have shown the bearing of the phenomena observed upon the constitution of the vapour in each case. In the present paper we propose to deal with chlorine, bromine, and iodine, the only other elements which lend themselves readily to this method of investigation.

Ten years ago the absorption spectra of bromine and iodine were investigated by Evans† by a method similar to that which we employ. He showed that the absorption varies with the temperature and pressure and drew from his results the conclusion that at sufficiently high temperatures the vapour of these elements is diactinic. This he explained by assuming that the monatomic molecules *into* which the more complex molecules are gradually resolved as the temperature is raised are colourless and without absorbent action on light at any rate within the limits λ 3500 to λ 6800. In most of his experiments the pressure was kept constant by connecting the observation tube in which the vapour was heated with a reservoir of the element maintained at a constant and lower temperature. The mass of the vapour at different temperatures was, therefore, variable. Our experiments were designed to show the effect produced by a constant mass of vapour on the absorption when the temperature was varied over as wide a range as possible. For this purpose a known quantity of the element under examination was enclosed in a sealed vacuous tube and the temperature gradually raised, observations upon the absorption being made at intervals. Under these conditions while the mass of the vapour remains the same throughout, the pressure increases both with the rise of temperature and as a consequence of the dissociation of the molecules. As, however, the tubes were evacuated with the Sprengel pump and the initial pressures were, therefore, low, it is improbable that the changes of pressure had any great influence upon the phenomena. This view is confirmed by the fact that observations made upon different quantities of the element showed only such differences of spectra as might be expected from the variation in the mass of the substance employed. In no case, so far as we observed, was the temperature of maximum absorption altered.

---

* 'Roy. Soc. Proc.,' A, vol. 95, p. 484 (1919); vol. 98, p. 147 (1920).

† 'Astrophysical Journal,' vol. 32, sec. 1, p. 291 (1910).

In our experiments on bromine, quantities varying from 1·4 mgrm. to 17 mgrm. were employed. The bromine was introduced into the observation tube by weighing a quantity of the dried double bromide of gold and potassium in a small capillary tube and placing this in the side tube* of the observation apparatus, evacuating, and sealing. The double bromide was then decomposed by heating, and the side tube sealed off at a point between the capillary tube and the body of the apparatus thus leaving bromine alone in the apparatus. By this means we were able without difficulty to get any required quantity of pure bromine into the tube. The quantity of bromine was checked by titration after each experiment. The observation tubes were of the same dimensions as those employed in the previous investigations (*loc. cit.*).

When the light from a Nernst filament is allowed to fall upon a white screen after passing through the vapour of bromine, a remarkable series of changes is observed as the temperature is raised. At 100° C. the colour, which is deep orange at lower temperatures, changes to brick red; above 600° C. the intensity of the transmitted light diminishes, the effect being most pronounced at 900° C. when the vapour is almost opaque. Above this point the brick-red colour changes to orange red; at 1200° C. it is pale yellow and at still higher temperatures nearly white.† In this connection it is interesting to recall that Andrews pointed out long ago‡ that when bromine is heated in a sealed tube it gradually becomes quite opaque.

Fig. 1 shows the absorption spectra of 9·7 mgrm. of bromine between the temperatures of 22° C. and 1350° C. It will be noticed that the absorption increases regularly up to about 900° C. after which it falls off continuously. The results are similar when larger quantities of bromine are employed (fig. 2), the absorption being of course greater, but the curve remaining of the same general form with the point of maximum absorption at the same temperature.

With 3 mgrm. of bromine a wide absorption band makes its appearance, the middle of which is situated about λ 4170. This band dies out about 600° C. but the spectrum remains weak in the same region until a much higher temperature is reached. Evans found in his experiments that there was no absorption at 1320° C., the highest temperature he employed. It seems possible, having regard to his method of observation, that what really happened was that the quantity of vapour remaining in his tube at the higher temperatures was too small to show any absorptive effect. In any case, his explanation that the disappearance of the absorption was due to the resolution of the diatomic into colourless monatomic molecules can hardly be correct seeing that at the highest temperature he reached it is probable that half the molecules remain undissociated.

---

* See fig. 1, 'Roy. Soc. Proc.,' A, vol. 95, p. 484 (1919).

† These observations were made on quantities of 8 or 9 mgrm. of bromine in observation tubes, 100 mm. long, 12·5 mm. diameter. When a larger quantity of bromine was employed, the light transmitted at 900° C. was so feeble that it could hardly be observed on the screen.

‡ 'Report Brit. Assoc., Edinb.,' 1871.

Iodine vapour, like that of bromine, changes colour as the temperature is raised. It was observed that with 4 mgrm., the colour which was reddish violet up to about 300° C. gradually deepened in shade to about 500° C. At this temperature the intensity of the transmitted light was greatly diminished and remained feeble up to about 800° C. when it again increased, the violet shade becoming lighter. At still higher temperatures the violet passed gradually into a bright salmon colour and finally the colour disappeared almost entirely.

The general character of the spectrum* is the same as in the case of bromine, but the point of maximum absorption, about 600° C., is much lower and the wide absorption band is much more distinctly defined. It will be remembered that the temperature of maximum absorption of tellurium is much higher than that of sulphur or selenium, the elements of lower atomic weight in the same group. With the halogens the reverse is the case. Figs. 3 and 4 show the absorption obtained with 2·8 and 4 mgrm. of iodine respectively. Fig. 4 shows a series of faint bands extending from λ 3990 to λ 4420 which first show distinctly at a temperature of 900° C. As the iodine was specially purified these bands must be due to the element and not to some foreign substance.

To obtain the spectrum shown in fig. 4 the temperature was raised to 1350° C., the highest point possible with silica tubes. It is interesting to observe that even at this temperature the wide band has not entirely died out, and the faint bands between λ 3990 and λ 4420 are still observable.† We are therefore unable to confirm Evans' statement as to the disappearance of the absorption spectrum at high temperatures. This is not surprising seeing that he examined the spectra by the eye. Without the aid of photography the faint lines would easily escape notice.

---

* Wood ('Phil. Mag.,' 1912, p. 680, and 1918, p. 236) has shown that the iodine absorption spectrum has about 50,000 bands in the visible.

† Spectrum in fig. 4 was taken with a plate more sensitive to the red than that employed in photographing the other spectra.

The investigation of chlorine presents peculiar difficulties. A large quantity of the element, involving the use of a long tube, must be employed to give a mass of the gas sufficient to exhibit the absorption phenomena. The spectrum shown in fig. 5 was obtained with a tube 43 cm. long containing chlorine at a pressure of 0·72 atmosphere at 0° C. It will be seen that there is a gradual increase in the amount of absorption with the rise of temperature, but at 1190° C., the highest temperature we reached in this case, no reversal such as is exhibited by bromine and iodine is apparent. Whether or not reversal would take place at a still higher temperature it is impossible to say with certainty. It may, however, be pointed out that the narrow bands at the red end of the spectrum have become much sharper at the highest temperature reached. Similar behaviour is observed in all those elements where the absorption is reversed, when the temperature of maximum absorption is being approached. The spectrum, which had been examined by Mrs. Laird to ascertain the effect of pressure at room temperature[*] is of the same fluted character as the spectra of bromine and iodine. At the higher temperatures faint diffuse bands are found in the region λ 5450 to λ 6180.

The vapour densities of the halogen elements show remarkable abnormalities. At low temperatures they are all somewhat higher than is required by the diatomic molecule. Thus, according to Treadwell and Christie,[†] the molecular weight of chlorine at 20° C. is 72. Pier,[‡] states that chlorine does not obey the ordinary gas laws below 184° C., but that between 300° C. and 1450° C. it is normal. According to Langer and V. Meyer,[§] the density of chlorine begins to fall off above 1200° C., the diminution being due to the breaking up of the diatomic molecules.

Jahn[||] states that the density of bromine does not become normal till over 200° C.[¶] Perman and Atkinson[**] show that about 900° C. the vapour is dissociated to a small extent. According to Victor Meyer,[††] the density of iodine vapour at 253° C. is a little higher than corresponds with the formula I_2. The dissociation of diatomic iodine molecules takes place, according to all the observers, at a lower temperature than in the case of bromine, but marked dissociation appears to begin between 600° C. and

---

[*] 'Astrophysical Journal,' vol. 14, p. 85 (1901).

[†] 'Zeit. für Angew. Chemie,' vol. 18, p. 1934 (1905).

[‡] 'Zeit. Physikal. Chem.,' vol. 62, p. 385 (1908).

[§] 'Ber.,' vol. 15, p. 2769 (1882).

[||] 'Gmelin Kraut.,' vol. 1, sec. 2, p. 226.

[¶] See also Perman, 'Roy. Soc. Proc.,' vol. 48, p. 45 (1890).

[**] 'Roy. Soc. Proc.,' vol. 66, p. 10 (1899).

[††] 'Ber.,' vol. 13, p. 394 (1880).

700° C.; 50 per cent. of the molecules are probably dissociated about 1270° C.*

It will thus be seen that in the case of the halogens there is evidence from the density that some of the molecules at low temperatures contain more than two atoms, and that at high temperatures the diatomic are mixed with monatomic molecules. It is remarkable that the temperatures at which the vapour density of bromine and iodine is stated to show a distinct falling off from the normal are precisely the temperatures at which we find the absorption to be at its maximum. Thus, according to Meier and Crafts,† the density of iodine begins to fall off between 600° C. and 700° C., and this is the temperature of maximum absorption. Again, according to Perman,‡ at 900° C. marked dissociation of bromine vapour is first observed; this is also the temperature of maximum absorption.

The fact that the absorption increases with the temperature up to a certain point is probably connected, as in the case of sulphur, with the gradual breaking down of more complex into simpler molecules. The vapour densities of the halogen elements at low temperatures point to the existence of aggregates of greater complexity than diatomic. As the temperature is raised these gradually dissociate, until the vapour consists mainly, if not entirely, of diatomic molecules, some of which in turn break up into monatomic molecules. It may be supposed that during this last phase the nascent atoms of the halogen tend to combine with one another and with the undissociated diatomic molecules to form more complex and more highly absorptive molecules; in other words, that the element polymerises. Applying this view to iodine vapour, the number of such polymers of iodine is at a maximum somewhat above 600° C.

In considering the relation of absorption to the molecular condition of the vapour, it is to be remembered that the vapour density is an average value for all the molecular aggregates present in the vapour at the temperature at which the density is taken. A constant value may occur for the density over a wide range of temperature while the absorption varies considerably, if we assume, for example, in the case of the halogens, that a sufficient number of molecules greater than X_2 are present to balance the diminution of density caused by the dissociation of X_2 to 2X. Thus, with iodine there is a range of several hundred degrees over which the density appears to remain constant while the absorption shows a marked increase. The spectroscopic method, therefore, indicates a change

---

* Naumann, 'Ber.,' vol. 13, p. 1050 (1880).

† 'Ber.,' vol. 13, p. 851 (1880).

‡ 'Roy. Soc. Proc.,' vol. 66, pp. 17, 489 (1900).

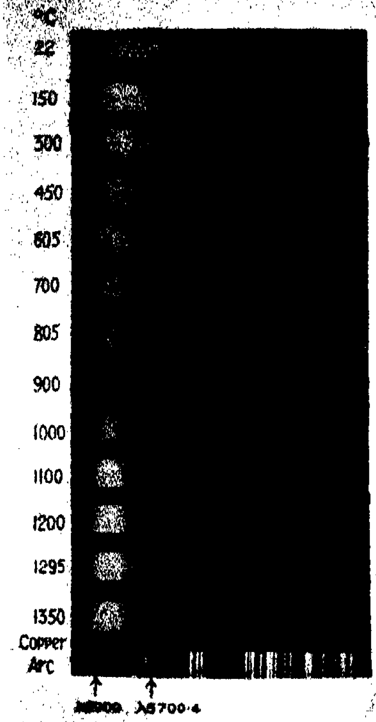


Fig 1.-Spectrum of 9·7 mgm of Bromine

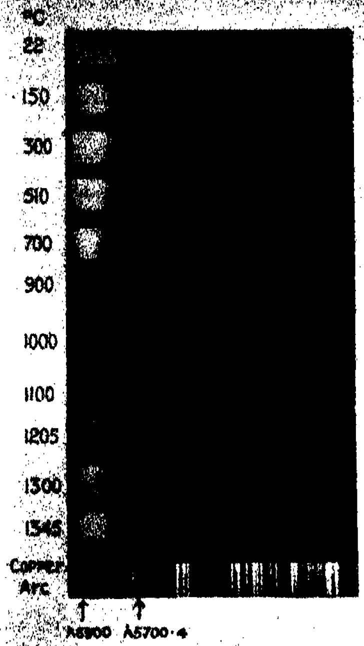


Fig 2.-Spectrum of 16·3 mgrm of Bromine

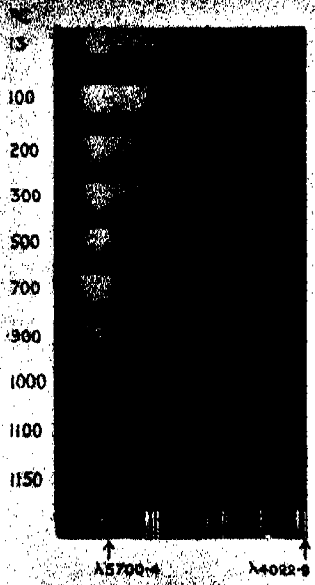


Fig 5.-Spectrum of Chlorine at 0·72 atmosphere initial pressure

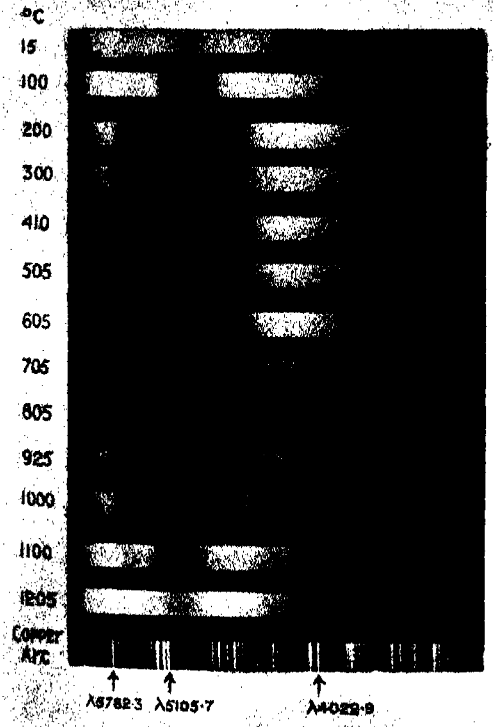


Fig 3.-Spectrum of 2·8 mgm of Iodine

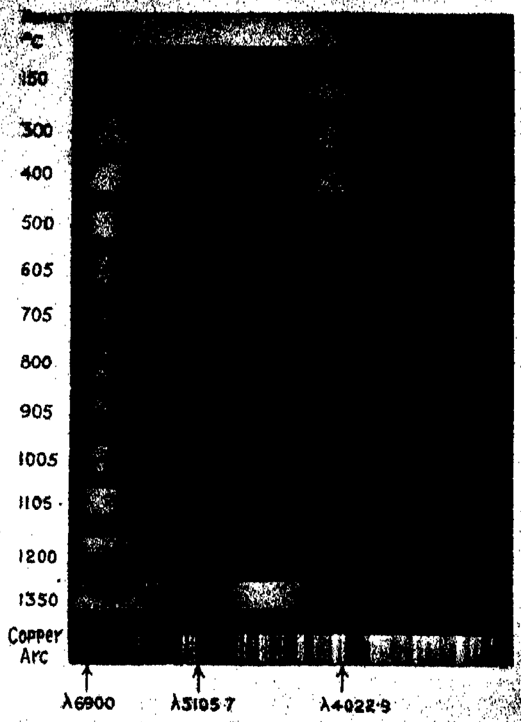


Fig 4.-Spectrum of 4·0 mgm of Iodine

in the constitution of the vapour upon which the density throws no light.

As in the case of sulphur, it is equally impossible to explain the changes in absorptive power of the halogens on the hypothesis that the diatomic molecules are gradually broken down into monatomic molecules, and that no other change takes place. This would leave the change which occurs at the temperature of maximum absorption unexplained. Wood and Kimura had already come to the conclusion on other grounds that the change is not so simple as is generally assumed. They point out* that, "on the assumption that diatomic absorbing iodine breaks down into a colourless monatomic gas we should expect the spectrum to fade away precisely as it does when the amount of vapour is decreased by lowering the density." By comparing the absorption exercised by a bulb containing iodine at high temperature (about 1000° C.) and at a temperature of 35° C. they found that some of the lines were much stronger in the spectrum of the cold vapour than in the spectrum of the hot vapour, while with others the reverse phenomenon was observed. They concluded that, inasmuch as the lines are affected in different degrees by an elevation of temperature, something more complicated takes place than the dissociation of diatomic into monatomic molecules.

The association of the temperature of maximum absorption with a tendency to polymerize is not only in accord with what is known as to the varying valencies of the halogens, but also brings the phenomena into line with those exhibited by sulphur which are almost certainly connected with the similar tendency of this element to undergo polymerisation.

---

* 'Astrophysical Journal,' vol. 46, p. 181 (1917).

PLATE 7.

The fine bands, though distinct on the negatives, are not clear in Plate 7, owing to the limitations of half-tone reproduction.

A Simple Extension of Fourier's Integral Theorem and some Physical Applications, in particular to the Theory of Quanta.

By R. H. Fowler.

(Communicated by Prof. Sir E. Rutherford, F.R.S. Received June 8, 1921.)

(1) *Introductory.*—In Poincaré's[*] proof of the necessity of Planck's hypothesis of quanta, an essential stage of the argument[†] depends on the use of Fourier's integral theorem to invert a particular infinite integral. In the form used by Poincaré this theorem may be enunciated thus :—

Under suitable conditions, if

$$\Phi(\alpha) = \int_0^\infty e^{-\alpha\eta} w(\eta)\, d\eta, \tag{1}$$

then

$$w(\eta) = \frac{1}{2\pi i}\int_C \Phi(\alpha)\, e^{\alpha\eta}\, d\eta, \tag{2}$$

where C *is a contour in the complex* α*-plane on which*[‡] $\mathbf{R}(\alpha) > \gamma > 0$ *and* $\mathbf{I}(\alpha)$ *goes from* $-i\infty$ *to* $+i\infty$.

Poincaré develops an argument which shows that, if $w(\eta)\,d\eta$ is the *a priori* probability that the energy of a resonating electron lies between η and $\eta + d\eta$, then $\Phi(\alpha)$ is such that $-d\{\log\Phi(\alpha)\}/d\alpha$ is the mean energy of the resonator at an absolute temperature C/α, where C is a known constant. When the mean energy of the resonator (of frequency ν) is known by experiment as a function of the absolute temperature, then $\Phi(\alpha)$ is known, except for an arbitrary constant multiplier. A direct appeal to formula (2) then shows that in these conditions, and with the same exception, $w(\eta)$ is also known and is, in fact, unique. It follows at once, and this is the object of Poincaré's work, that the known facts can be accounted for by one, and only one, function, $w(\eta)$—that is, in short, by the hypothesis of quanta.[§]

Unfortunately, the functions $w(\eta)$ and $\Phi(\alpha)$ that actually occur in the argument are such that Fourier's integral theorem does not apply at all. The actual value of $\Phi(\alpha)$ that turns up is $1/(1-e^{-\alpha\tau})$ and the integral (2) is meaningless. This leaves a gap in Poincaré's argument—a gap which it is desirable to fill. In order to do so, it is clear that we have to extend Fourier's integral theorem in such a way that an essentially discontinuous function such as the $w(\eta)$ of the last footnote is properly catered for. The type of extension required is to replace the integral (1) by what is really a Stieltjes' integral* of the form

---

[*] 'Journal de Physique,' Ser. V, vol. 2, p. 5 (1912).

[†] *Loc. cit.*, Sec. 7.

[‡] $\mathbf{R}(a)$ is the real part of a, $\mathbf{I}(a)$ is the real part of $-ia$.

[§] Planck's (first) hypothesis of quanta is, of course, that the energy of the resonator can only be of the form $p\epsilon$, where p is a positive integer or zero, and $\epsilon = h\nu$, where h is Planck's constant. Thus Planck's hypothesis attributes the value zero to $w(\eta)$ except when $\eta = p\epsilon$, when $w(\eta)$ becomes infinite in such a way that $\int_{p\epsilon-\delta}^{p\epsilon+\delta} w(\eta)\,d\eta = 1$. Poincaré's thesis is that this is the only possible form of $w(\eta)$ which will lead to the correct $\Phi(a)$ fitting experimental facts.

$$\int_0^\infty e^{-\alpha\eta}\, d\omega(\eta),$$

where $\omega(\eta)$ is a monotonic (increasing) function of η, and to obtain the analogue of Fourier's integral theorem in this case.

I am not aware whether or no any such extension of Fourier's integral theorem is possible for the general Stieltjes' integral indicated above, but for the purpose of physical applications it is unnecessary to contemplate an integral so general. We proceed in the following section to make a limited extension of Fourier's integral theorem in this direction, which is yet sufficient to enable us to complete properly this gap in Poincaré's argument. There appears to be a possibility of further applications of this extension which is noted in the final sections.

(2) *A Generalisation of Fourier's Integral Theorem.*—For physical applications we may and shall confine ourselves to monotonic functions $\omega(\eta)$ which have a finite number of steps only in any finite range (η_0, η_1). In such a case we may suppose that the steps occur at the values η_i of η $(i = 0, 1, 2, \ldots, \eta, \ldots)$ where $\eta_i \longrightarrow \infty$ steadily as $i \longrightarrow \infty$, and that the size of the step at η_i is j_i. In such a case we need not really introduce the fundamental conception of a Stieltjes' integral at all. We may *define* generally the integral $\int f(\eta) d\omega(\eta)$, over any finite range in which $f(\eta)$ is continuous, to be the sum of the series $\Sigma f(\eta_i) j_i$ and the more ordinary integral $\int f(\eta) d\omega'(\eta)$, where $\omega'(\eta)$ is a *continuous* monotonic (increasing) function of η, which is simply $\omega(\eta)$ with the steps cut out. We shall still further limit the generality of the function $\omega(\eta)$ by supposing that $\omega'(\eta)$ has a definite differential coefficient $\Omega\,(\geqq 0)$ which exists and is continuous at all except a finite of points in any finite range (η_0, η_1), and finally that Ω is bounded in any (η_0, η_1). Thus the integral $\int f(\eta) d\omega'(\eta)$ is merely the ordinary Riemann integral $\int f(\eta)\, \Omega\, d\eta$. This definition is in agreement, in the restricted cases to which it applies, with the usual direct definition of a Stieltjes' integral. We can extend the definition, both converge. We can also extend the definition to include discontinuities

---

* See *e.g.*, Hardy, 'Mess. Math.,' vol. 48, p. 90 (1919), where other references will be found.

to infinite ranges in the usual way provided that $\overset{\infty}{\Sigma} f(\eta_i) j_i$ and $\int^{\infty} f(\eta)\Omega d\eta$ and infinities of $f(\eta)$ provided that there are only a finite number of them in any finite range (η_0, η_1), that none of them coincides with any value of η_i, and, in the case of an infinity, that the corresponding part of the Riemann integral converges.

With these conventions as to the meaning of $\int f(\eta)\, d\omega(\eta)$ whenever such an integral occurs, we can proceed to our generalisation of Fourier's integral theorem. It will be well to enumerate precisely the conditions imposed by our conventions on $\omega(\eta)$. They are that *$\omega(\eta)$ is a monotonic (increasing) function of η for all values of η considered, which has only a finite number of simple discontinuities (steps) in any finite interval (η_0, η_1). The function $\omega'(\eta)$ which is derived from $\omega(\eta)$ by removal of the steps has a differential coefficient, Ω, which exists and is continuous except perhaps at a finite number of points in any finite range (η_0, η_1), and is also bounded in (η_0, η_1).* We shall refer to these conditions as conditions W. We can now enunciate and prove our extension of Fourier's integral theorem.

Theorem I.—The Extended Fourier's Integral Theorem.*

If $\omega(\eta)$ satisfies conditions W, *if $\Sigma j_i e^{-\alpha\eta_i}$ and $\int^{\infty} e^{-\alpha\eta}\, \Omega\, d\eta$ converge for $\alpha = \gamma_0$, and if*

$$\Phi(\alpha) = \int_{-a}^{\infty} e^{-\alpha\eta}\, d\omega(\eta), \tag{3}$$

then $\Phi(\alpha)$ is a holomorphic function of α in the half plane $\mathbf{R}(\alpha) > \gamma_0$, and

$$\tfrac{1}{2}\{\omega(\eta+0)+\omega(\eta-0)\} - \omega(-a) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} \Phi(\alpha)\, e^{\alpha\eta}\, \frac{d\alpha}{\alpha}. \tag{4}$$

where $\gamma > \gamma_0$, $\gamma > 0$, and the infinite integral in (4) *is to be evaluated as* $\underset{T\to\infty}{\mathrm{Lt}} \int_{\gamma-iT}^{\gamma+iT}$.

It is convenient in the enunciation of this theorem to use $-a$ and not zero as the lower limit of integration in (3), owing to the possibility of having the first step in $\omega(\eta)$ at $\eta = 0$. We can assume that a is so chosen that there is no step between $-a$ and 0.

---

* [Added, June 28.—In the limited case which we consider here, this extension of Fourier's integral theorem is practically equivalent to a combination of Fourier's ordinary integral theorem (§ (1)) with Perron's formula for the sum of the first n coefficients of a Dirichlet's Series (Hardy and Riesz, 'Dirichlet's Series,' Cambridge Tracts No. 18, p. 12). It is simpler and more satisfactory to give a direct proof of this extension, but this essential equivalence should not be overlooked. I am indebted to the referees for this footnote.]

Under the conditions of the theorem $\sum\limits_{i=0}^{\infty} j_i e^{-\alpha\eta_i}$ converges for $\alpha = \gamma_0$, and therefore represents a holomorphic function of α in the half plane $\mathbf{R}(\alpha) > \gamma_0$.* Under the same conditions $\int_{-a}^{\infty} e^{-\alpha\eta}\Omega d\eta$ also represents a holomorphic function of α in the same half plane.† Thus $\Phi(\alpha)$ is a holomorphic function of α in the region stated.

Now let

$$\begin{aligned} \mathrm{I}(\eta, \mathrm{T}) &= \frac{1}{2\pi i}\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} \Phi(\alpha)\, e^{\alpha\eta}\,\frac{d\alpha}{\alpha}, \\ &= \frac{1}{2\pi i}\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha\eta}\,\frac{d\alpha}{\alpha}\int_{-a}^{\infty} e^{-\alpha\zeta}\, d\omega(\zeta), \qquad (\gamma > \gamma_0). \end{aligned}$$

As $\gamma > \gamma_0$, the inner integral is by definition the sum of a uniformly (and absolutely) convergent series and a uniformly and absolutely convergent Riemann integral for all values of α on the path of integration. We can, therefore, invert the order of the two integrations‡ and get

$$\mathrm{I}(\eta, \mathrm{T}) = \frac{1}{2\pi i}\int_{-a}^{\infty} d\omega(\zeta)\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha(\eta-\zeta)}\,\frac{d\alpha}{\alpha}. \tag{5}$$

Now we know§ that if $\gamma > 0, \mathrm{T} > 0$, and $y > 1$,

$$\left| \int_{\gamma-\mathrm{T}i}^{\gamma+\mathrm{T}i} y^s \frac{ds}{s} - 2\pi i \right| \leq \frac{2}{\mathrm{T}}\frac{y^\gamma}{\log y}, \tag{6}$$

while if $y < 1$,

$$\left| \int_{\gamma-\mathrm{T}i}^{\gamma+\mathrm{T}i} y^s \frac{ds}{s} \right| \leq \frac{2}{\mathrm{T}}\frac{y^\gamma}{\log(1/y)}, \tag{6'}$$

and if $y = 1$,

$$\left| \int_{\gamma-\mathrm{T}i}^{\gamma+\mathrm{T}i} \frac{ds}{s} - \pi i \right| \leq \frac{2\gamma}{\mathrm{T}}. \tag{6''}$$

In our case $y = e^{\eta-\zeta}$. We therefore substitute from equations (6) in equation (5) and obtain

$$\begin{aligned} \mathrm{I}(\eta, \mathrm{T}) &= \frac{1}{2\pi i}\left\{\int_{-a}^{\eta-\epsilon} + \int_{\eta-\epsilon}^{\eta+\epsilon} + \int_{\eta+\epsilon}^{\infty}\right\} d\omega(\zeta)\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha(\eta-\zeta)}\,\frac{d\alpha}{\alpha}, \\ &= \int_{-a}^{\eta-\epsilon} d\omega(\zeta) + \underline{\mathrm{O}}\left\{\frac{1}{\mathrm{T}}\int_{-a}^{\eta-\epsilon}\frac{e^{\gamma(\eta-\zeta)}}{\eta-\zeta}\, d\omega(\zeta)\right\} \\ &\quad + \underline{\mathrm{O}}\left\{\frac{1}{\mathrm{T}}\int_{\eta+\epsilon}^{\infty}\frac{e^{\gamma(\eta-\zeta)}}{\zeta-\eta}\, d\omega(\zeta)\right\} + \frac{1}{2\pi i}\int_{\eta-\epsilon}^{\eta+\epsilon} d\omega(\zeta)\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha(\eta-\zeta)}\,\frac{d\alpha}{\alpha}. \end{aligned}$$

---

* Hardy and Riesz, *loc. cit.*, Theorem 4.

† Bromwich, 'Infinite Series,' p. 438.

‡ Bromwich, *loc. cit.*, pp. 117, 437.

§ Landau, 'Primzahlen,' vol. 1, p. 342.

Suppose η has any fixed value and that ϵ is chosen so small that the range $(\eta-\epsilon, \eta+\epsilon)$ contains no discontinuity of $\omega(\eta)$ other than (perhaps) η itself. Then

$$\mathrm{I}(\eta, \mathrm{T}) = \int_{-a}^{\eta-\epsilon} d\omega(\zeta) + \underline{\mathrm{O}}\left(\frac{1}{\mathrm{T}\epsilon}\right) + \frac{1}{2\pi i}\int_{\eta-\epsilon}^{\eta+\epsilon} d\omega(\zeta) \int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha(\eta-\zeta)}\frac{d\alpha}{\alpha}. \tag{7}$$

It is easily verified that, when $-\epsilon \leq \eta-\zeta \leq \epsilon$ (T fixed),

$$\frac{1}{2\pi i}\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha(\eta-\zeta)}\frac{d\alpha}{\alpha} = \underline{\mathrm{O}}(1),$$

and the integral is a continuous function of $\eta-\zeta$. Hence, if $\eta \neq \eta_i$, the last term in equation (7) is $\underline{\mathrm{O}}(\epsilon)$ and we have

$$\mathrm{I}(\eta, \mathrm{T}) = \omega(\eta-\epsilon) - \omega(-a) + \underline{\mathrm{O}}(1/\mathrm{T}\epsilon)\underline{\mathrm{O}} + (\epsilon).$$

Now let $\epsilon = 1/\sqrt{\mathrm{T}}$, and let $\mathrm{T} \to \infty$. We find

$$\underset{\mathrm{T}\to\infty}{\mathrm{Lt}}\ \mathrm{I}(\eta, \mathrm{T}) = \omega(\eta-0) - \omega(-a) = \omega(\eta) - \omega(-a).$$

If however $\eta = \eta_i$, then by definition

$$\frac{1}{2\pi i}\int_{\eta-\epsilon}^{\eta+\epsilon} d\omega(\zeta)\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha(\eta-\zeta)}\frac{d\alpha}{\alpha} = \frac{1}{2\pi i} j_i [\pi i + \underline{\mathrm{O}}(1/\mathrm{T})]$$
$$+ \frac{1}{2\pi i}\int_{\eta-\epsilon}^{\eta+\epsilon} d\omega'(\zeta)\int_{\gamma-i\mathrm{T}}^{\gamma+i\mathrm{T}} e^{\alpha(\eta-\zeta)}\frac{d\alpha}{\alpha} = \tfrac{1}{2}j_i + \underline{\mathrm{O}}(1/\mathrm{T}) + \underline{\mathrm{O}}(\epsilon).$$

Thus in this case

$$\underset{\mathrm{T}\to\infty}{\mathrm{Lt}}\ \mathrm{I}(\eta, \mathrm{T}) = \omega(\eta-0) - \omega(-a) + \tfrac{1}{2}j_i = \tfrac{1}{2}\{\omega(\eta-0) + \omega(\eta+0)\} - \omega(-a).$$

The proof of Theorem I is completed.

(3) *The Extended Mellin's Integral Theorem.*—The connection between Mellin's and Fourier's integral inversion formulæ is well known—they are in fact almost the same theorem.* It is, therefore, to be expected that a generalisation of Mellin's formula is possible similar to the above for Fourier's, and this generalization is easily made. We have

Theorem II.—The Extended Mellin's Integral Theorem.

If $\omega(x)$ *satisfies conditions* W, *if* $\Sigma j_i x_i^{-s}$ *and* $\int^{\infty} x^{-s}\Omega\, dx$ *converge for* $s = \gamma_0$, *and if*

$$\Phi(s) = \int_a^{\infty} x^{-s}\, d\omega(x),$$

---

* See *e.g.*, Hardy, 'Mess. Math.,' vol. 47, p. 178 (1918). [Added, June 28:—Formulæ (1) and (2) of this paper constitute Fourier's inversion formula (or integral theorem); the corresponding inversion formula of Mellin can be stated thus:—

Under suitable conditions, if

$$\phi(s) = \int_0^{\infty} w(x)\, x^{s-1} dx, \quad \text{then } w(x) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} \phi(s)\, x^{-s} ds.$$

The connection between the two formulæ may be seen by substituting $x = e^{\eta}$.]

where $a > 0$ *and* x^{-s} *has its principal value* $\exp(-s \log x)$, *then* $\Phi(s)$ *is a holomorphic function of* s *in the half plane* $\mathbf{R}(s) > \gamma_0$, *and*

$$\tfrac{1}{2}\{\omega(x+0)+\omega(x-0)\}-\omega(a) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\Phi(s)\, x^s \frac{ds}{s}, \tag{4'}$$

where $\gamma > \gamma_0$, $\gamma > 0$, *and the infinite integral in* (4′) *is to be evaluated as* $\operatorname*{Lt}_{T\to\infty}\int_{\gamma-iT}^{\gamma+iT}$.

We omit the proof which is identical with that of Theorem I. It is easy to extend the theorem to the case of $\int_0^\infty x^{-s}\, d\omega(x)$ provided that suitable convergency conditions at the lower limit are satisfied.

(4) *Application to the Theory of Quanta.*—Poincaré's argument adapted to our notation shows that if Planck's law for the mean energy of a resonator of frequency ν is true, and if

$$\Phi(\alpha) = \int_{-a}^{\infty} e^{-\alpha\eta}\, d\omega(\eta),$$

where $\omega(\eta)$ is the *a priori* probability that the resonator possesses an energy at any moment between 0 and η (limits included), then

$$\Phi(\alpha) = 1/(1-e^{-\alpha\epsilon}), \qquad (\epsilon = h\nu). \tag{8}$$

Equation (8) may be taken here as being demanded by the experimental facts. If we insert this value of $\Phi(\alpha)$ in equation (4) and apply Theorem I we find that the only value of $\omega(\eta)$ of the form stipulated by conditions W which can produce this $\Phi(\alpha)$ is given by the equation

$$\tfrac{1}{2}\{\omega(\eta+0)+\omega(\eta-0)\}-\omega(-\alpha) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{e^{\alpha\eta}}{1-e^{-\alpha\epsilon}}\frac{d\alpha}{\alpha}. \tag{9}$$

Suppose that η has any fixed value which is not equal to $p\epsilon$, where p is a positive integer or zero. For definiteness we may suppose that $p\epsilon < \eta < (p+1)\epsilon$. On expanding the integrand we obtain

$$\begin{aligned}\tfrac{1}{2}\{\omega(\eta+0)+\omega(\eta-0)\}-\omega(-a) &= \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\sum_{n=0}^{p} e^{\alpha(\eta-n\epsilon)}\frac{d\alpha}{\alpha} \\ &\qquad + \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{e^{-\alpha[(p+1)\epsilon-\eta]}}{1-e^{-\alpha\epsilon}}\frac{d\alpha}{\alpha}, \\ &= p+1+J.\end{aligned}$$

J is the value of the last term on the right. But since $(p+1)\epsilon-\eta > 0$, it is easily proved that $J = 0$ by applying Cauchy's theorem to the rectangle whose corners are $\gamma \pm i\mathrm{T}$, $\mathrm{X} \pm i\mathrm{T}$, and making $\mathrm{X} \longrightarrow \infty$, and $\mathrm{T} \longrightarrow \infty$. When

$\eta = p\epsilon$ the same reasoning applies except that the last term in the series contributes only $\frac{1}{2}$. Thus

$$\omega(\eta) - \omega(-a) = p+1, \qquad (p\epsilon < \eta < (p+1)\epsilon),$$
$$\tfrac{1}{2}\{\omega(\eta+0)+\omega(\eta-0)\} - \omega(-a) = p+\tfrac{1}{2}, \qquad (\eta = p\epsilon).$$

Now $\omega(\eta_1) - \omega(\eta_0)$ is the *a priori* probability* that the energy η of the resonator lies between η_0 and η_1. Hence this probability must be zero unless the range (η_0, η_1) includes one of the points $p\epsilon$, and unity when (η_0, η_1) includes any one such point. But this can only mean that the energy of the resonator may take any of the values $p\epsilon$ but no other values, and can therefore only change in either direction by jumps of amount ϵ, which is precisely Planck's (first) hypothesis. It hardly seems necessary to contemplate functions more general than those allowed by conditions W in a physical problem such as this.

In concluding this stage of the argument one should point out that, at one stage in Poincaré's proof, it is necessary to make use of the multiple integral.

$$\int_{-a}^{\infty}\dots\int_{-a}^{\infty} e^{-\alpha\Sigma\eta}\, d\omega(\eta_1)\dots d\omega(\eta_n)$$

which represents $[\Phi(\alpha)]^n$, and to show that it can be expressed in the form

$$\int_{-a}^{\infty} e^{-\alpha x}\, d\phi_n(x);$$

this integral by Theorem I gives rise to the relation

$$\tfrac{1}{2}\{\phi_n(x+0)+\phi_n(x-0)\} - \phi_n(-a) = \frac{1}{2\pi i}\int_c [\Phi(\alpha)]^n e^{\alpha x}\frac{d\alpha}{\alpha}.$$

There is little difficulty in the necessary modification of Poincaré's work, but it must be given at length if at all, and I shall not consider the details here.

Since these sections were written, I have found that the same point is referred to by Planck in the Poincaré memorial volume of the 'Acta Mathematica' which has just appeared.† In section (3) of this article Planck criticises Poincaré's conclusion that his (Planck's) second hypothesis of quanta—discontinuous emission and continuous absorption—is inadmissible. He points out (p. 394) the gap in Poincaré's argument that I have attempted to fill up in this paper and the need for such a more complete discussion. But he suggests that it may be due to this gap in the mathematics that Poincaré's argument distinguishes between the first and second hypotheses, allowing the first to be necessary and the second impossible. I do not think that Planck's contention can be upheld, so far at least as it applies to the possibility of error in this gap. If the mean energy of an oscillator is to be

---

* A constant multiplier omitted.

† 'Acta Mathematica,' vol. 38, p. 387 (1921).

$$\tfrac{1}{2}\epsilon+\frac{\epsilon}{e^{\alpha\epsilon}-1}$$

as Planck's second hypothesis requires, so that

$$\Phi(\alpha)=\frac{e^{-\frac{1}{2}\alpha\epsilon}}{1-e^{-\alpha\epsilon}},$$

and if the groundwork of Poincaré's argument is admitted to be correct, then it appears to follow inevitably that

$$\omega(\eta)=n, \qquad (n-\tfrac{1}{2})\epsilon<\eta<(n+\tfrac{1}{2})\epsilon.$$

But by the same reasoning that applies successfully to the former case, this must mean that the energy of the resonator can only be of the form $(n+\frac{1}{2})\epsilon$ and must therefore change discontinuously both in emission and absorption.

(5) *The Absorption of Non-homogeneous Beams of X-Rays.*—The foregoing ideas may possibly find another application in the discussion of absorption curves of non-homogeneous beams of X-rays, a question suggested to me by Mr. C. G. Darwin. Let μ be the absorption coefficient for rays of a definite wave-length, so that for such monochromatic rays

$$\mathrm{I}(x)=\mathrm{I}_0e^{-\mu x}. \tag{10}$$

In this equation I_0 is the original intensity of the beam, and $\mathrm{I}(x)$ the intensity after it has penetrated a distance, x, through the given absorbing medium. The general radiation from an X-ray bulb consists of radiation with a continuous distribution of intensity over a wide range of wave-lengths (or values of μ), and superimposed on this the characteristic radiation (a line spectrum) of the anticathode if it is excited. In this line spectrum the intensity is concentrated at given wave-lengths, or given values of μ. If, therefore, we attempt to define a function $\phi(\mu)$ such that $\phi(\mu)\,d\mu$ is the intensity in the original beam of the radiation whose absorption coefficient lies between μ and $\mu+d\mu$, we are unable to do so, since $\phi(\mu)$ becomes infinite when we reach a line of the characteristic spectrum. This is in fact just the case we have had to consider in the theory of quanta. To include it, we make use of a function $\psi(\mu)$ which may be discontinuous and is such that $d\psi(\mu)$ is the intensity between μ and $\mu+d\mu$. The intensity $\mathrm{I}(x)$ in the non-homogeneous beam after penetrating a thickness, x, will be given by the equation

$$\mathrm{I}(x)=\int_0^\infty e^{-\mu x}\,d\psi(\mu), \tag{11}$$

which is an integral of just the form we have been discussing in this paper. We may assume that $\psi(\mu)$ satisfies conditions W. The problem before us is to determine $\psi(\mu)$, the intensity distribution from the known values of $\mathrm{I}(x)$.

From the physical conditions of the problem $\int_0^\infty d\psi(\mu)$ must converge, for the intensity in the original beam is finite. The integral (11) therefore exists for all values of $\mathbf{R}(x) \geqq 0$ and represents a holomorphic function of x in the right-hand half-plane. On applying Theorem I we should find at once that

$$\tfrac{1}{2}\{\psi(\mu+0)+\psi(\mu-0)\}-\psi(0) = \frac{1}{2\pi i}\int_c \mathrm{I}(\alpha)\, e^{\alpha\mu}\,\frac{d\alpha}{\alpha},$$

where C is the contour $a-i\infty, a+i\infty, a > 0$. As it stands however this formula is probably of little or no practical value, for $\mathrm{I}(x)$ is a function of x which is defined (observationally) only for real values of x. It will not in general be possible to deduce directly from the observed values of $\mathrm{I}(x)$ its proper functional form or even an approximation to its proper functional form for general complex values of x. In fact it is only by a somewhat fortunate chance that this direct use of Theorem I can be made in the application to quanta. The function $\Phi(\alpha)$ there is really only determined (observationally) for real α, but it so happens that it has the precise functional form $1/(1-e^{-\alpha\epsilon})$ which defines it for complex α as well.

We can, however, turn the difficulty by the method applied by Schwartzschild* to a somewhat similar problem in the theory of stellar distributions. The method consists essentially in a use of Mellin's integral (Theorem II) to construct from $\mathrm{I}(x)$ a function $\mathrm{J}(q)$ which is known as a complex function of q—to which, therefore, Fourier's (or in this case Mellin's) integral formula can be applied.

Equation (11) defines $\mathrm{I}(x)$. Let

$$\mathrm{J}(q) = \int_0^\infty \mathrm{I}(x)\, x^{q-1}\, dx, \qquad [\mathbf{R}(q) > 0]. \tag{12}$$

The integral converges absolutely at the lower limit if $\mathbf{R}(q) > 0$. It converges at the upper limit under the physical conditions of the problem, for in any beam of radiation there is a minimum wave-length—a minimum value μ_0 of μ—below which the intensity is necessarily zero. Thus, in effect

$$\mathrm{I}(x) = \int_{\mu_0}^\infty e^{-\mu x}\, d\psi(\mu),$$

and it is evident that as $x \longrightarrow \infty$

$$\mathrm{I}(x) = \underline{O}(e^{-\mu_0 x}).$$

We find therefore that

$$\mathrm{J}(q) = \int_0^\infty x^{q-1}\, dx \int_{\mu_0}^\infty e^{-\mu x}\, d\psi(\mu).$$

---

* **Eddington, 'Stellar Motions and the Structure of the Universe,' p. 208.**

This is an absolutely convergent repeated integral and the order of integration may, therefore, be inverted* provided that the inverted integral exists, as, in fact, it does here. This gives

$$J(q) = \int_{\mu_0}^{\infty} d\psi(\mu) \int_0^{\infty} e^{-\mu x} x^{q-1}\, dx,$$

or

$$\frac{J(q)}{\Gamma(q)} = \int_{\mu_0}^{\infty} \mu^{-q}\, d\,\psi(\mu), \qquad [\mathbf{R}(q) > 0)]. \tag{13}$$

Applying Theorem II, we find at once that

$$\tfrac{1}{2}\{\psi(\mu+0)+\psi(\mu-0)\} = \frac{1}{2\pi i}\int_c \frac{J(q)}{\Gamma(q)}\,\mu^q \frac{dq}{q}, \qquad [\psi(\mu_0-0) = 0]. \tag{14}$$

Thus given the numerical values of the function $I(x)$ for real values of x, it is theoretically possible by two sets of integrations to determine the function $\psi(\mu)$ for all real values of μ. This method has actually been used in the problem of stellar distributions,† but it requires precise numerical data to make it practicable. Whether it could be applied to the analysis of actual X-ray absorption curves is open to question, and I do not propose to pursue the matter further here.

(6) *Summary.*—In this paper I have attempted to fill up a distinct gap in Poincaré's discussion of the necessity for quanta. It is at once clear that a mild form of Stieltjes' integral is a necessary instrument for this purpose; such integrals appear to occur naturally in other physical problems. The machinery of Fourier's integral theorem generalised in such a direction is thus of some importance; it is evident that the theorem must be true under much wider conditions than those under which I have proved it, and I recommend its more complete discussion to those who are interested in this branch of pure mathematics.

---

* The inner integral is, of course, the sum of an absolutely and uniformly convergent series and an absolutely convergent Riemann integral.

† Eddington, *loc. cit.*, p. 220.

Some Problems connected with Evaporation from Large Expanses of Water.

By M. A. GIBLETT, M.Sc.

(Communicated by Dr. G. C. Simpson, F.R.S. Received May 10, 1921.)

Introductory.

1. The water which falls as rain is originally evaporated either from land surfaces, forests, etc., or from the oceans, in particular from the latter. It is therefore of importance to ascertain how the evaporated water distributes itself in an atmospheric current during its progress across a water surface, and, further, to investigate the relation between the length of the path over the water surface and the amount of water thereby rendered available. In this paper, these problems are treated for the case of a current of air of uniform speed moving over a water surface of uniform temperature. An empirical formula is employed to represent the rate of evaporation from each element of the water surface, and account is taken of the stirring upwards of the evaporated water by the agency of turbulence, assuming the latter to be uniformly distributed throughout the current of air. The results obtained would find application, for example, in discussing problems connected with evaporation in the trade wind zones, or from inland seas and lakes, or, in particular, from the North Sea, winds from which frequently bring much cloud and quite appreciable rainfall to the eastern coasts of Great Britain and the northern coasts of France. All the constants occurring in the formulæ obtained are known, either accurately or approximately, and, accordingly, the order of magnitude of the evaporation from a given stretch of water, under assigned conditions, can be estimated. What is also of importance, is the comparison of the amount of water evaporated from a given area or of the distribution of water vapour in the air above it, under one set of conditions, with those under a different set. Of particular interest is the effect of varying the speed of the air.

2. The suggestion which led to the present investigation is contained in a paper on "The Meteorological Conditions of an Ice Sheet and their bearing on the Desiccation of the Globe."* The author of that paper, in considering the evaporation in tropical and sub-tropical regions during the Quaternary Ice Age as compared with that at the present time, has occasion to compare the rate of evaporation from *large* expanses of water under different

---

* C. E. P. Brooks, 'Quarterly J. Roy. Meteorolog. Soc.,' vol. 40 (1914).

conditions as to wind and surface temperature, and Lt.-Col. E. Gold, in the discussion, emphasises the fact that this is in reality a very complex problem for the formula used (quoted below, p. 474) for the rate of evaporation from an element of the surface, involves the vapour pressure at the dew point of the air "near" the surface. Now this vapour pressure varies from point to point along the current of air which is traversing the water, so that an integration is necessary in order to obtain the rate of evaporation from a large stretch of water, and, further, the vapour pressure at any point depends on the speed with which the air has reached that point, which complicates the problem of comparing the evaporation under different wind conditions.

3. Before proceeding further, it should be mentioned that the quantity described as the vapour pressure "near" the surface is not so indefinite as might at first sight appear to be the case. It is known that, though the air flowing over a water surface is in general in a turbulent condition, yet, next to the water, there is a thin layer, consisting of mixed air and water vapour, the molecules of which are only very slowly replaced by those from the turbulent region above, and which probably remains unbroken by waves, unless conditions are such as to produce spray. In this layer the diffusion is purely molecular, but, beyond it, turbulence endows the air with a greatly increased diffusivity for water vapour. The consequence is that the humidity in general decreases very rapidly from saturation actually at the water surface to some lower value at the point where the turbulent *régime* is entered, after which the change, with increasing distance from the surface, is comparatively slow. Jeffreys[*] estimates the thickness of this surface layer as of the order of 1 mm. or more, while Bigelow[†], in the account of his practical researches in evaporation, refers to a thin film, next to the water surface, in which the humidity passes rapidly from saturation to its value near the surface where the tubes of flow of water vapour are "distorted, broken, carried away in the wind, and converted into a heterogeneous mass of shattered tubes." He states that the gradient of vapour pressure is rapid in the first few millimetres and then slow up to great heights. Marvin,[‡] too, states that, close down to the free water surface there must be a thin layer of air, heavily charged with water vapour, which, from our knowledge of viscosity and the flow of fluids, is, in spite of moderate wind action, changed and renewed by the wind only with relative slowness. He points out that the dense vapour sheet must be

---

[*] H. Jeffreys, "Some Problems of Evaporation," 'Phil. Mag.,' p. 35, March, 1918.

[†] F. H. Bigelow, "Studies of the Phenomena of the Evaporation of Water over Lakes and Reservoirs," 'Monthly Weather Rev.,' 1907, 1908, 1910; also 'Bulletin of the Argentine Meteorological Office,' No. 2, Part I, March, 1911: Part II, July, 1912.

[‡] C. F. Marvin, "A Proposed New Formula," 'Monthly Weather Rev.,' February, 1909.

very thin, for observations at Reno[*] showed that the humidity one half-inch above the water surface differed little from the value two or three feet higher. It is the region just outside the thin superficial layer which is referred to as "near" the surface.

4. Reference to the annotated bibliography of evaporation prepared by Mrs. Grace J. Livingston[†] shows that a number of empirical formulæ have, from time to time, been given to represent the rate of evaporation from an exposed water surface. In evaluating the constants in these formulæ, observations of the evaporation from water in pans have generally been used, and there is abundant evidence to show that the results so obtained cannot be transferred to evaporation from the surface of a large expanse of water without correction. A few investigators have, however, taken account of this aspect of the problem, so that there are empirical formulæ available which are designed to apply to the evaporation from an element of a large water surface.

Of the various empirical formulæ the one commonly known as Dalton's has received a great deal of support. Its form is

$$E = A(P_W - p_0)(1 + cv),$$

where E is the rate of evaporation per unit area, P_W the maximum water vapour pressure corresponding to the temperature of the water surface, p_0 the actual water vapour pressure in the air "near" the surface, v the speed of the air "near" the surface, and A and c constants. The constants A and c have been evaluated by a number of investigators, but there is lack of agreement between the results. This is partly due to the use of evaporating pans of different sizes and exposure. Fitzgerald[‡] and Carpenter,[‡] however, working independently with various exposures, including tanks or pans floating in water, obtained values of A in very good agreement, but widely different values were obtained for the wind coefficient c. This may be partly explained by the fact that while Fitzgerald measured the wind near the surface of the water, Carpenter used records from an anemometer on the roof of a building away from the tanks. The most comprehensive series of experiments in this field of research is that of Bigelow (*loc. cit.*) who, however, was lead to a different formula.[§] Nevertheless he recognises Dalton's formula as approximate and

---

[*] F. H. Bigelow, "Studies of the Phenomena of the Evaporation of Water over Lakes and Reservoirs," 'Monthly Weather Rev.,' February, 1908.

[†] 'Monthly Weather Rev.,' June, September, November, 1908; February, March, April, May, June, 1909.

[‡] "Annotated Bibliography of Evaporation," Mrs. Grace J. Livingston, 'Monthly Weather Rev.,' March, 1909.

[§] Bigelow's formula is $$E = B \frac{P_w}{p_0} \left(\frac{dP}{dT}\right)_w (1 + cv),$$ where Pw, p_0, v and c, have the same significance as in Dalton's formula, B is a constant gives values for the constants in it, applicable to open water and based on his extensive observations in places with climates sufficiently different to ensure a wide range of meteorological conditions. The remaining formulæ need not be discussed in detail, for none, up to the present, seems to have been established as giving a better representation than the above, of the rate of evaporation from an element of the surface of a large expanse of water. Dalton's formula leads to a readily integrable case and will, therefore, be adopted. For numerical examples Bigelow's estimate of the coefficients in it must be accepted as the best available.

5. The problem is that of the upward diffusion of water vapour in a current of air of uniform speed, with a boundary condition at the surface given by the above empirical formula. The upward flux of water vapour per unit area at any level is represented by the expression $-\mathrm{K}\rho\,\partial\mu/\partial z$, where μ is the mass of water vapour per unit mass of air, z the height above the surface, ρ the density of the air and K the coefficient of eddy-diffusivity. Evidence is gradually being collected as to the values which K may assume, and measurements by Richardson[*] indicate an increase with height, at least in the first 1000 metres above the surface. However, in the present instance, only a first approximation will be made and K will be assumed independent of height z.

Notation.

The system is referred to rectangular axes with the origin at a fixed point "near" the water surface, and the z-axis vertical, while the air is supposed moving horizontally in the direction of the y-axis. The motion is turbulent, but the mean[†] conditions at any point (y, z) are supposed constant and the following notation is used to represent them:—

ρ, the density of the air.
μ, the mass of water vapour per unit mass of air.
W, the absolute humidity $(= \mu\rho)$.
H, the relative humidity.
T, the absolute temperature.
p, the actual pressure of water vapour.

---

and $\left(\frac{d\mathrm{P}}{d\mathrm{T}}\right)_{\mathrm{W}}$ is the rate of change of maximum water vapour pressure with temperature at the temperature of the water surface.

[*] L. F. Richardson, "Some Measurements of Atmospheric Turbulence," 'Phil. Trans.,' A, vol. 221 (1920).

[†] Time average over an interval centred at any particular instant and sufficiently long to smooth out the minor fluctuations due to turbulence. The turbulence is assumed uniformly distributed, so that the interval would be the same for all parts of the air current. A more rigorous definition hardly seems necessary here.

P, the maximum water vapour pressure at temperature T.

v, the velocity of the air,

K, the coefficient of eddy-diffusivity in a vertical direction.

t, the time taken by a sample of air to travel a distance y from origin ($= y/v$).

A suffix 0 denotes conditions at the level $z = 0$, while a dash denotes conditions at $y = 0$. T_W denotes the temperature of the water surface and P_W the corresponding maximum vapour pressure.

Mathematical Analysis.

6. Let D/Dt denote $\partial/\partial t + v\partial/\partial y$, then if the vertical flux of water vapour at any point is $-K\rho\,\partial\mu/\partial z$ and the horizontal flux $-K_y\rho\,\partial\mu/\partial y$, where K_y is used for the moment to denote the coefficient of eddy-diffusivity in a horizontal direction, the "equation of continuity" gives,

$$\frac{D}{Dt}(\rho\mu) = \frac{\partial}{\partial y}\left(K_y\rho\frac{\partial\mu}{\partial y}\right) + \frac{\partial}{\partial z}\left(K\rho\frac{\partial\mu}{\partial z}\right).$$

Now K_y may or may not be equal to K, but in any case the first member on the right is negligible in comparison with the second, except perhaps in the immediate vicinity of the origin $y = 0$, in consequence of the very slow horizontal variations of the meteorological elements in this problem as compared with the vertical variations. If, further, the state is steady, so that $\frac{\partial}{\partial t}(\rho\mu) = 0$, then

$$v\frac{\partial}{\partial y}(\rho\mu) = \frac{\partial}{\partial z}\left(K\rho\frac{\partial\mu}{\partial z}\right).$$

K is to be taken independent of z and thus, neglecting $\partial\rho/\partial z$

$$\frac{\partial\mu}{\partial y} = \frac{K}{v}\frac{\partial^2\mu}{\partial z^2}.$$

The rate of evaporation per unit area, $A(P_W - p_0)(1 + cv)$ must equal the value of $-K\rho\,\partial\mu/\partial z$, the vertical flux, at $z = 0$, so that the boundary condition to be satisfied there is

$$(-K\rho\,\partial\mu/\partial z)_0 = A(P_W - p_0)(1 + cv).$$

Now $p = \alpha TW = \alpha T\rho\mu$, where $\alpha = 4{\cdot}616 \times 10^6$ when C.G.S. units are used,* and thus the boundary condition becomes

$$(\partial\mu/\partial z)_0 = -h(1 - \beta\mu_0)/\beta,$$

where $$\beta = \alpha T_0\rho_0/P_W$$

and $$h = A\alpha T_0(1 + cv)/K,$$

---

* 'Computer's Handbook,' Meteorological Office, 223, Introduction. London, 1916.

We require, therefore, a solution of

$$\frac{\partial \mu}{\partial y} = \frac{K}{v}\frac{\partial^2 \mu}{\partial z_2}$$

with conditions $(\partial \mu/\partial z)_0 = -h(1-\beta\mu_0)/\beta$

and $\mu' = f(z),$

the latter denoting the arbitrary distribution of water substance above the origin.

Write $\Omega = 1-\beta\mu.$

Then $$\frac{\partial \Omega}{\partial y} = \frac{K}{v}\frac{\partial^2 \Omega}{\partial z^2}$$

with conditions $(\partial \Omega/\partial z)_0 = h\Omega_0$

and $\Omega' = f_1(z)$

where $f_1(z) = 1-\beta f(z).$

If v and T_0 are supposed independent of y and if vertical variations of v are neglected, the solution* is

$$\Omega = h\int_0^\infty \phi(z+\zeta, t)\, e^{-h\zeta}\, d\zeta,$$

where $$\phi(z, t) = \frac{1}{2\sqrt{(\pi K t)}}\int_0^\infty F(z')\left[e^{-(z-z')^2/4Kt} - e^{-(z+z')^2/4Kt}\right] dz',$$

$$F(z) = f_1(z) - f_1'(z)/h \qquad \text{and} \qquad t = y/v.$$

This solution holds for the arbitrary distribution of water vapour at $y = 0$ given by $\mu' = f(z)$. The expression may be converted into one containing known functions only if $f(z)$ is taken as a constant, which describes the case of air which has descended or passed off dry land where it has been well stirred up. Then introducing the notation $\sigma = zh/2$ and $\tau = h\sqrt{(Kt)}$, and omitting the analysis, the solution becomes

$$\Omega/\Omega' = \Theta(\sigma/\tau) + e^{-(\sigma/\tau)^2}\psi(\sigma/\tau+\tau),$$
$$= \chi(\sigma, \tau),$$

where $$\Theta(x) = \frac{2}{\sqrt{(\pi)}}\int_0^x e^{-u^2}\, du,$$

and $$\psi(x) = \frac{2}{\sqrt{(\pi)}}\int_x^\infty e^{x^2-u^2}\, du.$$

Note that $\psi(x) = \chi(0, x).$

---

* See Carslaw, 'Introduction to Fourier Series,' Macmillan.

Relative humidity may now be introduced if we note that

$$\Omega = 1-\beta\mu = 1-\frac{\alpha \mathrm{T}_0\rho_0\mu}{\mathrm{P_W}} = 1-\frac{\mathrm{PT}_0\rho_0}{\mathrm{P_W T}\rho}\cdot\frac{\alpha \mathrm{T}\rho\mu}{\mathrm{P}}$$

$$= 1-\frac{\mathrm{PT}_0\rho_0}{\mathrm{P_W T}\rho}\cdot\frac{p}{\mathrm{P}} = 1-\frac{\mathrm{H}}{\lambda},$$

where $$\lambda = \mathrm{P_W T}\rho/\mathrm{PT}_0\rho_0.$$

Also

$$\Omega' = 1-\beta\mu' = 1-\frac{\alpha \mathrm{T}_0\rho_0\mu'}{\mathrm{P_W}} = 1-\frac{\mathrm{P}_0}{\mathrm{P_W}}\frac{p_0'}{\mathrm{P}_0} = 1-\frac{\mathrm{P}_0}{\mathrm{P_W}}\mathrm{H}_0'.$$

While it is not essential, we shall take the surface air temperature equal to that of the surface water, so that $\mathrm{T}_0 = \mathrm{T_W}$ and $\mathrm{P}_0 = \mathrm{P_W}$.

Then $\Omega/\Omega' = (1-\mathrm{H}/\lambda)/(1-\mathrm{H}_0')$, but $\Omega/\Omega' = \chi(\sigma, \tau)$ and hence the relative humidity at any point is given by

$$\mathrm{H} = \lambda\{1-(1-\mathrm{H}_0')\chi(\sigma, \tau)\}.$$

If $z = 0$, then $\sigma = 0$ and $\lambda = 1$, and the expression for the relative humidity "near" the surface is

$$\mathrm{H}_0 = 1-(1-\mathrm{H}_0')\psi(\tau).$$

The rate of evaporation per unit area at any point of the surface, given by Dalton's formula, $\mathrm{A}(\mathrm{P_W}-p_0)(1+cv)$, may be written $\mathrm{AP}_0(1-\mathrm{H}_0)(1+cv)$, which on substituting for H_0 becomes $\mathrm{E}'\psi(\tau)$, where $\mathrm{E}' = \mathrm{AP}_0(1-\mathrm{H}_0')(1+cv)$, the rate of evaporation per unit area at the origin.

The mass of water evaporated per unit time from a strip of unit breadth, extending from the origin a distance y along the y-axis is accordingly given by

$$\mathrm{E}'\int_0^y \psi(\tau)\,dy \quad \text{where} \quad \tau = h\sqrt{(\mathrm{K}y/v)}.$$

Performing the integration, the result is

$$\mathrm{E}'y\nu(\tau) \quad \text{where} \quad \nu(\tau) = [2/\tau\sqrt{(\pi)}-\{1-\psi(\tau)\}/\tau^2].$$

The average rate over the strip, or the mean depth evaporated from the strip per unit time, is thus $\mathrm{E}'\nu(\tau)$. Again, the rate of evaporation from the strip of length y must equal the difference between the rate at which water is carried across a vertical plane normal to the current at distance y from the origin and the rate at which it is carried across a similar plane at the origin, since a steady state is considered and there is no accumulation of water vapour in any region. Hence, if all the water gained by a column of air on a base of unit area in passing a distance y over the water surface could be precipitated instantaneously, the depth of rain produced would be

$$\mathrm{E}'y\nu(\tau)/v \quad \text{or} \quad \mathrm{E}'t\nu(\tau).$$

7. Collecting the more important formulæ, we find that the relative humidity at any point, under the conditions laid down in the course of the preceding work is given by

$$H = \lambda\{1-(1-H_0')\chi(\sigma, \tau)\}, \tag{1}$$

where $\sigma = zh/2; \qquad \tau = h\sqrt{(Ky/v)},$

$$h = A\alpha T_0(1+cv)/K; \qquad \lambda = P_0T\rho/PT_0\rho_0.$$

"Near" the surface the relative humidity is,

$$H_0 = 1-(1-H_0')\psi(\tau). \tag{2}$$

The rate of evaporation, per unit area, at any point of the surface is

$$E'\psi(\tau), \tag{3}$$

while the mean depth of evaporation, per unit time, from a strip of length y, extending from the origin, is

$$E'\nu(\tau), \tag{4}$$

where E' is the rate of evaporation, per unit area, at the origin, and equals $AP_0(1-H_0')(1+cv)$.

Finally, the water gained by a column of air on a base of unit area in passing a distance y over a water surface, would, if it could all be precipitated instantaneously, produce a depth of rain equal to

$$E't\nu(\tau). \tag{5}$$

Fig. 1 shows the way in which $\chi(\sigma,\tau)$ depends on σ and τ. Each curve

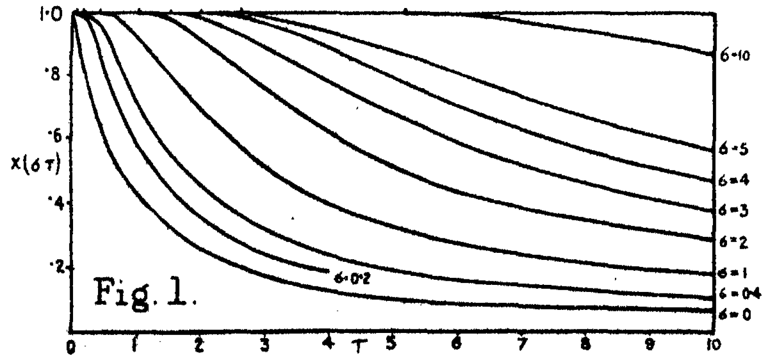
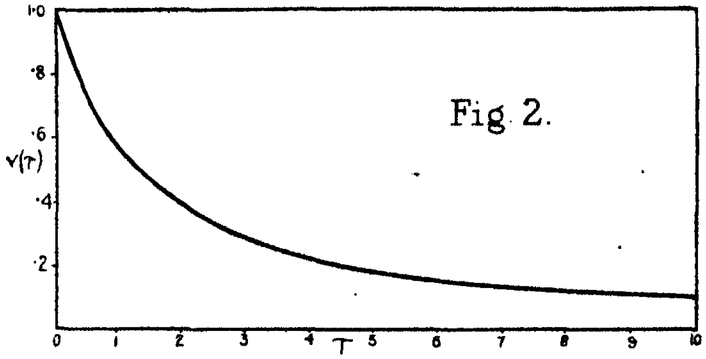


Fig. 1.

Representing the functions χ(σ, τ) and ψ(τ).

represents it as a function of τ for a given value of σ. The curve for $\sigma = 0$ represents the function $\psi(\tau)$. $\nu(\tau)$ is given in fig. 2.



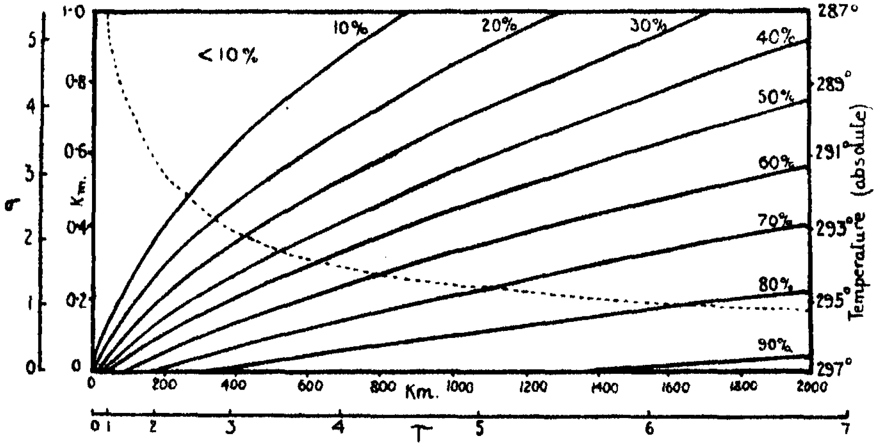
Representing the function V(τ).

Numerical Examples.

8. However, before discussing these results, it will be well to introduce some numerical values for the constants in them. Taking, as mentioned above, Bigelow's estimate (*loc. cit.*) for the constants in the evaporation formula, and converting them into C.G.S. units, we obtain $A = 3{\cdot}14 \times 10^{-10}$ and $c = 3{\cdot}024 \times 10^{-3}$, when E is measured in cm. sec.$^{-1}$, or g. sec.$^{-1}$, v in cm. sec.$^{-1}$ and the vapour pressures in dynes cm.$^{-2}$. These figures are for fresh water. For salt water it is thought that the evaporation should be somewhat less, but the correction is not known definitely and depends on the salinity. A further complication, of which the effect is unknown, is the formation of spray, which must, when forming, seriously disturb the restraining superficial vapour film referred to in an early paragraph. The value of K is not a constant of the atmosphere, and, indeed, may vary over a very large range. Richardson (*loc. cit.*), by observations on the scattering of smoke from a steamer, has found a value over the open sea as low as 10^2 C.G.S. units and as high as 10^4 C.G.S. units. Values much lower than 10^2 are likely to occur, but it is probable that 10^4 is approaching the upper limit for conditions over the open sea. In the examples given the value 10^4 is adopted. For the case of a wind speed of 5 m. sec.$^{-1}$, air initially perfectly dry, a surface temperature of 297° a. (75·2° F.), and a lapse rate of 1° C. per 100 m., the preceding analysis yields the distribution of water vapour shown in fig. 3. The wind is supposed blowing from the left to the right of the diagram, which represents a section 2000 kilom. in length and 1 kilom. in height. The full lines indicate surfaces of equal relative humidity, while the dotted line represents the height at which cloud would commence to form in a sample of air transferred from the surface directly upwards without loss or gain of heat. In fig. 4 the mass of water vapour per mass of air is supposed initially (*i.e.*, on the left of the diagram) constant at all heights, and such as to give a relative humidity of

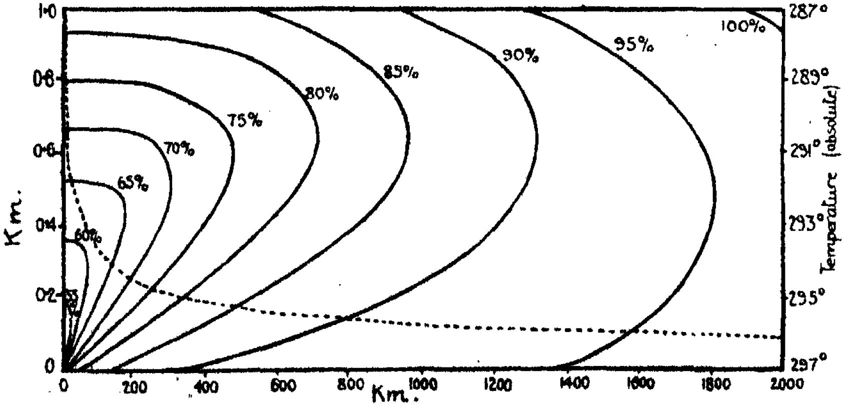
50 per cent. at the surface, while the remaining conditions are the same as in fig. 3. This initial condition implies cloud above the height of about 1350 m., and the diagram shows that this cloud gradually descends until at 2000 kilom. from the origin its base is in the neighbourhood of 1 kilom. above the surface, the 100 per cent. isopleth coming just below this level on

Fig. 3.



Showing the distribution of water vapour in a current of air, initially dry, flowing with uniform speed of 5 $m.sec^{-1}$ over a water surface of temperature 297°a (=75·2°F), the lapse rate of temperature being 1°C per 100 m and the coefficient of eddy-diffusivity 10^4 C.G.S. units

Fig. 4.



Showing the distribution of water vapour under the same conditions as in fig (3) but with the mass of water vapour per mass of air initially (i.e on the left of the diagram) the same at all heights and such as to produce a relative humidity of 50% at the surface.

the right of the diagram.* This is, of course, only an approximate representation, and no account is taken of the thermo-dynamical complications due to latent heat or of precipitation.

It should be noted, in connection with figs. 3 and 4, that detached cloud may occur even in regions where the mean relative humidity is less than 100 per cent. Where the mean relative humidity is high this is very likely. For example, the dotted line in fig. 4 shows that, at a distance of, say, 1600 kilom. from the origin, a sample of air near the surface has not to be projected to any great height, in the course of the turbulent flow, for condensation to occur. The introduction of τ and σ in place of y and z greatly simplified the construction of these diagrams and the necessary computations. The scales for τ and σ appropriate to the conditions of figs. 3 and 4 are shown in fig. 3.

9. The surface temperature in these two cases is applicable to trade wind regions, and an appropriate initial condition would probably be somewhere between the two, the air being less damp at the higher levels than in the second example. The case represented would be one in which the trade wind clouds are detached cumuli or entirely absent, for a continuous sheet of cloud would imply a mean condition of saturation. The dotted lines in figs. 3 and 4, while marking the condensation level for air rising directly upwards from the surface, would not necessarily mark the level of the bases of such cumuli, for they may be formed by condensation in a rising column of air formed by samples gathered from various levels. They set, however, in these examples, a lower limit to this level. Observations show that the clouds in the trade winds often assume the form of a continuous sheet of strato-cumulus, but this probably corresponds in general to a condition which is supposed not to hold in the present paper, namely, that the region affected by turbulence extends a limited distance only above the surface of the water. Such cloud sheets, often associated with inversions of temperature, and marking the upper limit of the turbulent region, have frequently been observed. Aviators generally find "smooth" flying above them, though the air may be very "bumpy" below.

10. The curves of fig. 5 show the rate of evaporation at each point of a water surface under various conditions as to wind, etc., while those of fig. 6 show the corresponding mean rates from strips of water extending from the origin any distance in the direction of the wind. The ordinate at any point in fig. 6 is accordingly equal to the mean ordinate, between that point and

---

* The 100 per cent. isopleth is to be distinguished from the dotted line in the diagram. The latter merely marks the condensation level for air rising directly upwards from the surface.

the origin, of the corresponding curve in fig. 5. Thus, under the conditions of curve 1, fig. 5 shows that the rate of evaporation at a point distant 1000 kilom. from the origin is about $8{\cdot}5 \times 10^2$ mm. year^{-1}, while fig. 6 shows that the mean rate from the whole strip, extending from the origin to this point, is about $1{\cdot}4 \times 10^3$ mm. year^{-1}.

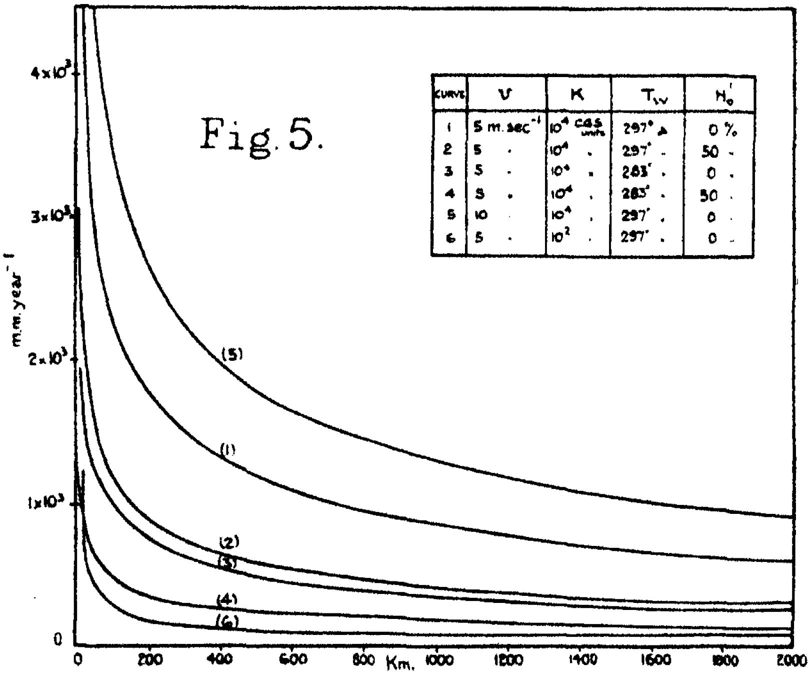


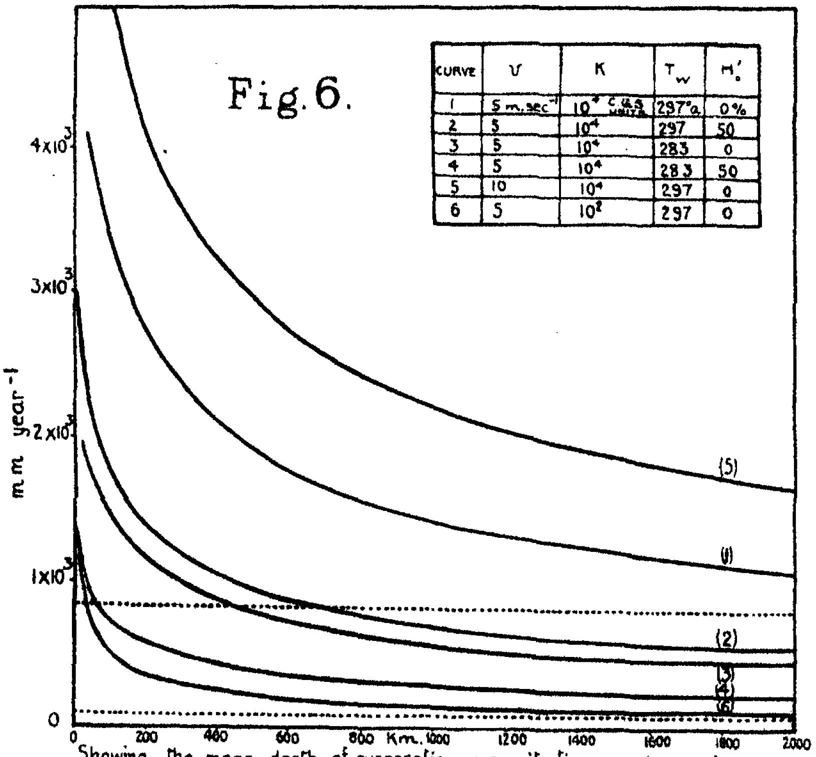
Fig. 5.

CURVE	V	K	T_w	H_0'
1	5 m. sec⁻¹	10⁴ c.g.s. units	297° a.	0 %
2	5 .	10⁴ .	297° .	50 .
3	5 .	10⁴ .	283° .	0 .
4	5 .	10⁴ .	283° .	50 .
5	10 .	10⁴ .	297° .	0 .
6	5 .	10² .	297° .	0 .

Showing the rate of evaporation at each point of a water surface under various conditions as to wind etc. [computed from expression (3)]. The wind is supposed blowing from the left to the right of the diagram.

Curves 1 and 2 correspond respectively to the conditions of figs. 3 and 4, and reference to expressions 3 and 4 shows that curve 2 represents a rate of evaporation one half that represented by curve 1. In order to illustrate the influence of surface temperature on the rate of evaporation, curves 3 and 4 have been drawn, corresponding to the same conditions as the first two curves, but with surface temperature 283° a. (50° F.) instead of 297° a. (75·2° F.). Such a surface temperature would apply, at some period of the year, to the North Sea, and fig. 6 shows that, under the conditions of curve 3, the mean depth of water evaporated from a stretch 1000 kilom. in length would be of the order of 6×10^2 mm. year^{-1}. Under the conditions of curve 4, it would be only half this. Utilising expression 5, it is found that, under the conditions of curve 3, the air, after traversing 1000 kilom. of the North Sea, would have gained enough water vapour to produce, if precipitated instantaneously, a depth of rain of the order of 4 mm. Again, this would be halved under the conditions of curve 4. Curve 5 replaces curve 1 if the wind speed is changed from 5 m. sec.$^{-1}$ (11·2 miles hour^{-1}) to 10 m. sec.$^{-1}$, while curve 6 replaces curve 1 if the value of K is reduced to 10^2. The dependence of the rate of evaporation from a body of water on the value of K is well illustrated by this example.

Fig. 6.



CURVE	V	K	T_w	H'_0
1	5 m. sec^{-1}	10^4 C.G.S. units	297°a.	0%
2	5	10^4	297	50
3	5	10^4	283	0
4	5	10^4	283	50
5	10	10^4	297	0
6	5	10^2	297	0

Showing the mean depth of evaporation per unit time, under various conditions, from stretches of water extending any distance in the direction of the wind, [computed from expression (4)].

11. Murray* has estimated the mean annual rainfall on the land of the globe as equivalent to 8·4 × 10^2 mm. (upper dotted line in fig. 6), and that only a fraction 1/4·5 of this, finds its way to the oceans by rivers. This fraction spread over the oceans amounts† to 7·4 × 10 mm. (lower dotted line in fig. 6), and must be the excess of evaporation above precipitation over the oceans, if we assume that the mean annual evaporation over the whole globe equals the mean annual precipitation. The fact that the dotted lines repre-

---

* J. Murray, 'Scottish Geographical Magazine,' February, 1887.

† Taking the ratio of land surface to water surface as 1·45 to 3·67. See 'Computer's Handbook,' Meteorological Office, 223, Introduction. London, 1916.

senting these quantities in fig. 6 fall where they do, when plotted on the same scale as the curves representing the rates of evaporation from large stretches of ocean, under selected sets of possible conditions, is of great interest, and shows, at least, that the numerical quantities employed in these examples are of the right order of magnitude.

Some General Conclusions.

12. Let us now compare the distribution of water vapour in the air over a water surface, and also the rate of evaporation from it, under different conditions as to wind, surface temperature and turbulence.

Consider, firstly, the relative humidity "near" the surface as given by expression (2).

$$H_0 = 1-(1-H_0')\,\psi(\tau) \quad \text{where} \quad \tau = A\alpha T_0(1+cv)\sqrt{(y/Kv)}.$$

Since $\psi(\tau)$ is a function which decreases as τ increases (see fig. 1), it is evident that H_0, the relative humidity "near" the surface must increase as τ increases. Other quantities remaining constant, τ increases as y increases, hence the relative humidity "near" the surface increases with increased distance from the origin, as is to be expected. This is illustrated in figs. 3 and 4. Again, considering conditions at an assigned distance y from the origin, a greater value of T_0 implies a greater value of τ, so that a higher surface temperature implies a higher relative humidity. It is to be noted, however, that τ is only altered in the ratio of the absolute temperatures, so that the increase in relative humidity at a given point "near" the surface, corresponding to a higher surface temperature, is not very great for the ordinary range of sea surface temperatures which occur in nature. Thus, conditions being the same as in fig 3, the relative humidity at a distance of 171 kilom. from the origin, corresponding to a value of τ equal to 2, is 74·5 per cent., while if the surface temperature were 283° a., instead of 297° a., the value of τ would become 1·91 and the relative humidity consequently 73·5 per cent., a difference of only 1 per cent. The rate of evaporation from the strip of water extending from the origin 171 kilom. down the wind would, however, be very different in the two cases, as is seen by referring to curves 1 and 3 of fig. 6. In fact, it is the greater evaporation, accompanying the higher surface temperature, which leaves the relative humidity at an assigned distance from the origin, very little changed.

As regards turbulence, it is seen that a decrease in the value of K implies an increase in τ and a corresponding increase in the value of the relative humidity at a given point near the surface. The effect, too, may be large, for as has been indicated, K may vary over a large range. A rather large value,

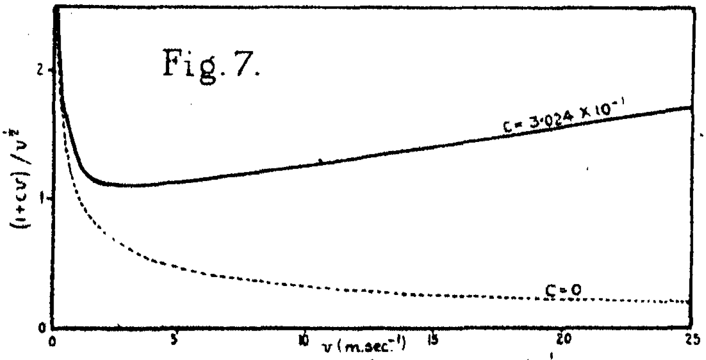
10^4 C.G.S. units, was adopted for K in the examples of figs. 3 and 4. Suppose K decreased to 10^3 C.G.S. units, then τ is multiplied by 10, so that the point in fig. 3, distant 171 kilom. from the origin, which there corresponds to $\tau = 2$ now corresponds to $\tau = 20$, and the relative humidity increases from 74·5 to 97·2 per cent. The effect on the rate of evaporation from strips of various lengths is seen by comparing curves 1 and 6 in fig. 6.

The speed of the air v enters the expression for τ in the form $(1+cv)/v^{\frac{1}{2}}$. This function of v is plotted in fig. 7 for Bigelow's value of c and for the value $c = 0$. When $c = 0$ it is seen to decrease as v increases, so that τ decreases and consequently the relative humidity at a given point "near" the surface decreases. The effect of c, which is positive, is, however, to counteract this and, if Bigelow's value is correct, to reverse it for wind speeds above about 3 m. sec.$^{-1}$, an increase in the relative humidity at a given distance from the origin, corresponding to an increased wind speed. It is here appropriate to emphasise the twofold manner in which the wind speed enters into the problem. If c were zero, then, as has been seen, the greater the wind speed the less would be the relative humidity H_0 at every point "near" the surface and hence, by reference to the evaporation formula which may then be written $E = AP_0(1-H_0)$, the greater the evaporation. In this effect the wind speed merely operates by controlling the rate at which drier air is carried along to replace that rendered humid by evaporation and account is taken of it in the mathematical analysis. However, all experiments, such as those of Bigelow and Fitzgerald, designed to study the rate of evaporation from an element of a large water surface point to a value of c, far different from zero. Thus at a given instant, at any point of a water surface, for a given value of the surface temperature and of the vapour pressure at the dew point of the air "near" the surface, a greater rate of evaporation corresponds to a higher wind speed "near" the surface. This is an experimental result and the factor $(1+cv)$ in the evaporation formula must be regarded as taking account, empirically, of the effect of the wind speed "near" the surface on the process of the diffusion of water vapour between the actual surface of the water and the point defined earlier as "near" the surface.

So far each of the conditions affecting evaporation has been varied independently of the others. The effect of an increase in K, corresponding to an increase in v, should be considered though there are no grounds for supposing that K can be expressed directly as a function of v. Suppose, purely for the sake of illustration, that K is proportional to v, then $\tau \propto (1+cv)/v$, *i.e.*, $\tau \propto 1/v+c$, and consequently an increase in v definitely implies a decrease in τ and hence a decrease in the relative humidity "near" the surface at a given distance from the origin.

13. Similarly, the effect of changes in the conditions, on the humidity at heights above the surface, may be readily found from expression 1, through the medium of τ and σ, in conjunction with fig. 1. One point worthy of notice is that, for a given value of σ, the function $\chi(\sigma, \tau)$ does not depart appreciably from unity until τ reaches a value given by $\tau = \sigma/2$, as is at once evident from fig. 1. This means that, at the height corresponding to σ, the humidity does not begin to be affected by the rising water vapour until the air reaches a distance from the origin corresponding to a value of τ equal to $\sigma/2$. In the example of fig. 4, for instance, the humidity at height 740 m. ($\sigma = 4$) is first affected when the air has travelled a distance of about 171 kilom. ($\tau = 2$) from land. The relation $\tau = \sigma/2$ is equivalent to $z = 4\sqrt{(\mathrm{K}y/v)} = 4\sqrt{(\mathrm{K}t)}$, and does not involve the surface temperature T_0. In his investigation of the formation of fog over the Newfoundland Banks, Taylor[*] found that, when air flows over colder water, the height to which cooling extends is given by $z = 2\sqrt{(\mathrm{K}t)}$ in the present notation. This height appears to disagree with the above result by being only half as great. Taylor, however, was concerned with the thickness of the "inversion" of temperature, and reference to the diagram on p. 35 of the 'Manual of Meteorology'[†] shows that cooling commences to be appreciable at twice this height, or where $z = 4\sqrt{(\mathrm{K}t)}$ in agreement with the present result.

Fig. 7.



Representing the function $(1+cv)/v^{\frac{1}{2}}$

14. The effect of changes in the conditions, on the mean depth of evaporation from strips of water extending from the origin any distance in the direction of the wind, has been considered incidentally in the preceding remarks. It may be studied by considering expression 4 in relation to fig. 2. As has been seen, Bigelow's estimate of the coefficient c implies that τ increases slowly as v increases, for values of v above about

---

[*] G. I. Taylor, "On Eddy-Motion in the Atmosphere," 'Phil. Trans.,' A, vol. 215, p. 1.

[†] Sir Napier Shaw, 'Manual of Meteorology,' Part IV. Camb. Univ. Press, 1919.

3 m. sec.$^{-1}$. Hence $\nu(\tau)$ decreases (see fig. 2). However, it is evident from the manner in which v enters the expression for E′, that, in spite of this, the rate of evaporation, $E'\nu(\tau)$, increases as v increases. This is exemplified by curves 1 and 5 of fig. 6. Exactly the same argument applies to the case of surface temperature, and the greater evaporation accompanying a greater surface temperature is shown by curves 1 and 2 as compared with 3 and 4 in fig. 6. The coefficient of eddy-diffusivity (K) enters the expression for the rate of evaporation solely through τ, and, since τ decreases as K increases, this is accompanied by an increased rate of evaporation. Since K may vary over a wide range, its effect on the evaporation may be very great. An example is furnished by curves 1 and 6 of fig. 6.

15. These last remarks furnish examples of Brooks's problem, which originated this enquiry. It was, as mentioned earlier, the comparison of the evaporation from a large expanse of water of uniform surface temperature, under different conditions as to wind and temperature. He made a rough comparison by applying Dalton's formula to the whole area, assuming H_0, the relative humidity of the air "near" the surface, to have the same value everywhere, and to be the same in the two cases compared. In this paper it has been shown how H_0 varies in the direction of the wind, and that, under the conditions laid down, the mean depth of evaporation per unit time, from a long stretch of water in this direction, is given by the formula (4),

$$AP_0(1-H_0')(1+cv)\nu(\tau).$$

This amounts to applying Dalton's formula to the water surface at the origin, and then multiplying by the factor $\nu(\tau)$, which depends on the wind speed, water surface temperature, eddy-diffusivity, and the length of the stretch of water. However, adopting Bigelow's value of c, it has been seen that $\nu(\tau)$ varies only slightly in comparison with the other factors, when the wind speed and water surface temperature are altered, so that a rough comparison of the rates of evaporation from the strip may be made by comparing the rates of evaporation at the origin, as given by Dalton's formula. This, of course, supposes that the eddy-diffusivity, which may affect $\nu(\tau)$ considerably, remains unaltered.

The actual cases compared by Brooks are (1), $v = 20$ m.p.h. (8·9 m. sec.$^{-1}$), $T_0 = 80°$ F. (299·7° a.), and (2) $v = 25$ m.p.h. (11·2 m. sec.$^{-1}$), $T_0 = 75°$ F.) (296·9° a.), the relative humidity at all points being assumed to be 50 per cent. If we suppose the relative humidity of 50 per cent. to apply to the origin and K and y to be the same in both cases, and distinguish between the cases by suffixes 1 and 2, then

$$\tau_2/\tau_1 = 1{\cdot}05 \quad \text{since} \quad \tau = A\alpha T_0(1+cv)\sqrt{(y/Kv)}.$$

If R denote the rate of evaporation from a stretch of water extending from the origin down wind, then from (4)

$$R_2/R_1 = 1{\cdot}01 \times \nu(\tau_2)/\nu(\tau_1).$$

Now if $\tau_2 = \gamma\tau_1$, where γ is a constant greater than unity, then it may be shown that, as τ_1 increases from 0 to ∞, $\nu(\tau_2)/\nu(\tau_1)$ decreases from unity to $1/\gamma$, whatever the value of K may be. In the present instance $\gamma = 1{\cdot}05$, so that as the length of the stretch of water is increased from 0 to ∞, R_2/R_1 decreases from 1·01 to 0·96, and this is independent of the value of K, provided it is the same in both cases. Thus, whatever the length of the strip, the evaporation is very little different in the two cases—the conclusion reached by Brooks.

16. It has been shown that the rate of evaporation, $E'\nu(\tau)$, from a stretch of water increases with the wind speed. Nevertheless, the water gained by a column of air in passing a given distance over the water surface, decreases as the wind speed increases. For a column of air, on base of unit area, the quantity referred to is, from (5), $E't\nu(\tau)$, which may be written $AP_0(1-H'_0)(1/v+c)y\nu(\tau)$. Now, for values of v greater than 3 m. sec.$^{-1}$, $\nu(\tau)$ has been shown to decrease as v increases, and the other factor involving v clearly does so, with the result stated.

Summary.

17. Air passing over a water surface is generally turbulent, but in the immediate vicinity of the surface is a thin superficial layer, which mixes only slowly with the air above. In this layer the humidity passes rapidly from saturation actually at the water surface to a lower value at the base of the turbulent region, after which the change with height is slow. The region where the lapse rate of humidity first becomes small is referred to as "near" the surface. To take account of the molecular processes at the water surface and in the superficial layer of air in contact with it, an empirical formula is chosen, expressing the rate of evaporation per unit area in terms of the temperature of the water surface and the vapour pressure and speed of the air "near" the surface. Taking this formula as expressing the boundary condition at the surface, the problem of the upward stirring of water vapour in a current of air, of uniform speed, moving over a water surface of uniform temperature, is dealt with by mathematical analysis, and the resulting distribution of water vapour and rate of evaporation from large stretches of water, obtained. The turbulence is supposed uniformly distributed throughout the current of air above the level defined as "near" the surface, the coefficient of eddy-diffusivity in this region being taken as independent of

height. Certain approximations are made, in that the rate of change with height of the air speed and density are neglected.

Some particular cases are worked out, but in view of the approximations made and some uncertainty as to the value of the constants in the empirical evaporation formula, and indeed as to its form, the results give an indication of the order of magnitude, rather than precise numerical estimates, when applied to natural conditions.

Finally, the effect of the wind speed, the water surface temperature, and the eddy-diffusivity, on the evaporation from a large stretch of water and the distribution of water vapour above it, is examined by varying each individually within the natural range of values. The effect of eddy-diffusivity in particular, which may vary through a large range, is found to be very pronounced.

I wish to take this opportunity of expressing my thanks to Lt.-Col. E. Gold, F.R.S., both for suggesting this problem and for his helpful criticism and advice during the course of the work.

---

*A Study of Catalytic Actions at Solid Surfaces.** VI.—*Surface Area and Specific Nature of a Catalyst: two Independent Factors controlling the Resultant Activity.*

By E. F. Armstrong, D.Sc., F.R.S., and T. P. Hilditch, D.Sc.

(Received May 12, 1921.)

Catalytic action is determined by two primary factors, the presence of a mechanically suitable surface at which the process takes place, and the simultaneous presence of the chemical agent (catalyst) by which the change is promoted.

The work of Langmuir,† Bancroft,‡ Taylor§ and others upon the phenomena of adsorption has emphasised anew the importance of surface with reference to catalysis. When a suitable surface is present, however, no catalytic action will occur unless on that surface (whether it be formed by an inert support or by the catalyst alone) there is a specific chemical able to form appropriate intermediate compounds with the interactants: for example, if hydrogen and ethylene are in contact with such a surface, no formation of ethane will occur unless there is present some specific material, such as palladium, platinum, nickel or copper.

---

* Part V, 'Roy. Soc. Proc.,' A, vol. 98, pp. 27–40 (1920).

† 'J. Amer. Chem. Soc.,' vol. 38, p. 2221 (1916); vol. 39, p. 1848 (1917); vol. 40, p. 1361 (1918).

‡ Presidential Address, American Electrochemical Society, 1920.

§ 'J. Ind. Eng. Chem.,' vol. 13, p. 75 (1921).

It is probable that confusion sometimes has occurred because the respective functions of general surface and specific chemical have not been sufficiently understood to be differentiated, and an account of some work with catalytic nickel, which illustrates somewhat concisely the incidence of each factor, may therefore be of interest.

The first group of experiments correlate the apparent volume occupied by (or bulk gravity of) various catalysts with their activity; next, the rate of reduction of nickel oxide is shown to be similarly connected with bulk gravity, and finally we give an example of the relation between the degree of reduction of nickel oxide deposited upon a suitable support and the activity of the reduced catalyst in oil-hydrogenation.

Apparent Volumes of various types of Nickel Oxide and the corresponding Reduced Metals.

Kelber* showed that, whilst nickel oxide *per se*, when reduced at about 300° C., gave a moderately active catalyst, reduction of the same material at 450° C. gave a product of but feeble activity; yet nickel oxide deposited upon kieselguhr, and reduced at 450°, gave a catalyst much more active than the unsupported metal prepared at 300° C.

In the Table below are given data which compare the apparent volumes of various oxides of nickel and their reduction products with their densities, pyrophoric character and catalytic activity.

The apparent volumes were measured by shaking down the substances in narrow stoppered graduated tubes, and are given as the volume (number of cubic centimetres) occupied by 1 grm. of the material.

The samples studied were:—

(i) *Powdered fused nickel oxide*, prepared by igniting finely-divided nickel in a current of oxygen, and grinding the cooled mass of fused oxide.

(ii) *Nickel hydroxide*, prepared from the sulphate or nitrate by precipitation, washing and cautious drying.

(iii) *Nickel hydroxide on kieselguhr*, prepared as in the last case, but in presence of kieselguhr.

---

* 'Ber.,' vol. 49, p. 55, 1868 (1916).

Source of reduced nickel.				Reduced nickel.			
	Density.	Apparent volume.	Temperature of reduction.	Density.	Apparent volume.	Pyrophoric character.	Catalytic activity.
		c.c.	° C.		c.c.		
Powdered fused oxide	6·96	0·85	500	8·14	0·52	None	None
Precipitated hydroxide	5·41	0·87	300 500	7·85 8·18	0·83 0·56	Pronounced Medium	Fair Very little
On kieselguhr...	1·63	3·22	500	1·85	2·67	None	Very active

Attention is drawn to the apparent volumes of the reduced precipitated hydroxide: at 300° C. this is almost the same at that of the unreduced substance, but considerable shrinkage has occurred during reduction at the higher temperature of 500° C.

The fused oxide, on the other hand, occupies less than half the bulk of the precipitated hydroxide, although its specific gravity is not much greater; and, during reduction, a certain amount of expansion (increase in porosity) is noticeable, but the product is quite inactive.

On the other hand, in the supported nickel, the kieselguhr, which does not shrink on heating, almost wholly determines the bulk of the product, and the activity is correspondingly great.

The column showing the pyrophoric character of the reduced metals is added only because it illustrates that the parallelism between this property and catalytic activity, at one time supposed to be complete, is in reality by no means so: the most active catalyst of the series is not pyrophoric, or at least only to a very small degree.

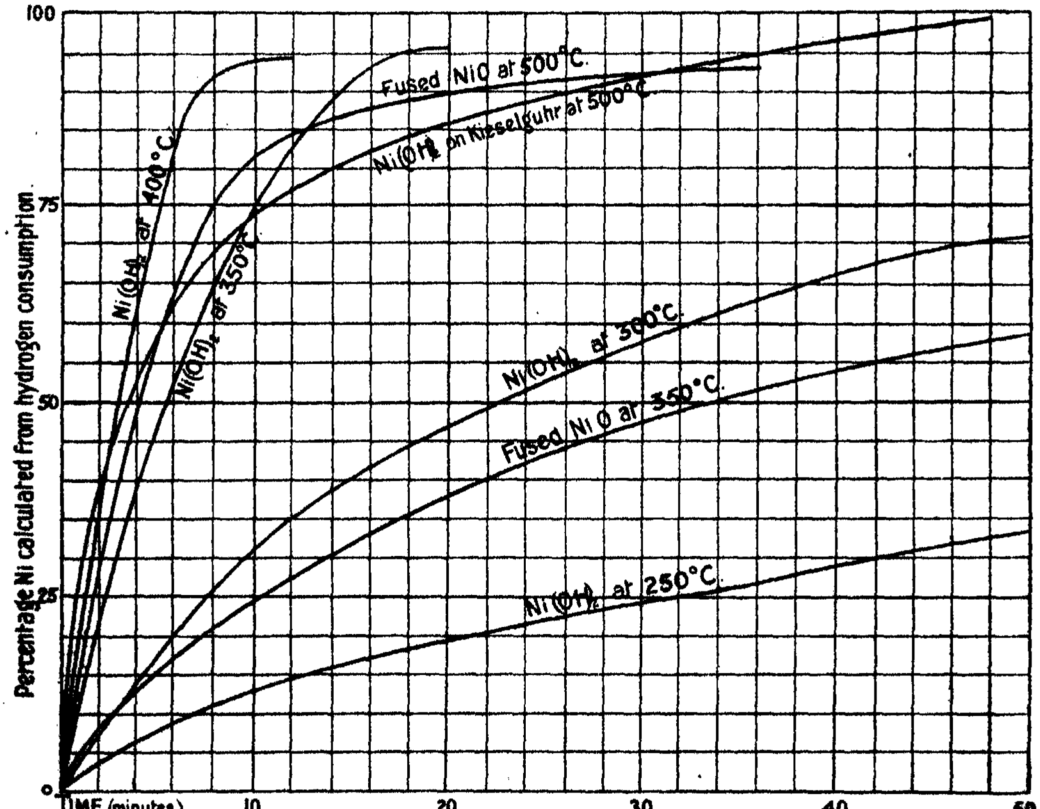
Rate of Reduction of various types of Nickel Oxide.

It has frequently been observed, with unsupported nickel oxide, that a partially reduced oxide is more active than the same when completely reduced, and this fact has been made much of by those workers who have supposed that the active agent in hydrogenation is not nickel, but a sub-oxide of the metal. Sabatier and Espil* showed, however, that, although a point of inflexion occurs in the curve showing the amount of reduction of nickel oxide (at 175° C.) against time at a point corresponding to an oxide between Ni_3O and Ni_4O, there is evidence of the presence of metallic nickel before 10 per cent. of the oxide has been reduced, and these authorities are of the opinion that the amount of nickel then present, though small, is amply sufficient to account for the hydrogenating power of the product.

---

* 'Compt. rend.,' vol. 158, p. 674 (1914).

We reproduce here a series of curves (hydrogen consumed by reduction of nickel oxide/time) obtained in the reduction of the oxides quoted in the above Table, from which it appears that the point of inflexion is mainly determined by the physical state of the oxide.



Rate of reduction of Nickel Oxides at varying temperatures.

We consider that the more rapid initial reduction represents the production of metallic nickel from the surface of the oxide particles, the slower subsequent rate corresponding to the reduction of a denser central core of oxide. Since, as our bulk gravity figures show, there is a tendency to contraction when a light precipitated oxide is reduced, it is obvious that a partially reduced oxide will consist of a film of nickel dispersed over the central core of each particle, and it is practically certain that the superior activity of this product over the fully reduced metal is due to its being a "supported" catalyst (nickel distributed on nickel oxide) and therefore of the same type, although not of the same degree of activity, as nickel upon a support such as kieselguhr.

The graphs shown in the figure present the reduction of precipitated nickel hydroxide at 250°, 300°, 350° and 400° C., of the fused nickel oxide at 350° and 500° C., and of nickel hydroxide on kieselguhr at 500° C.

The curve for precipitated nickel hydroxide reduced at 250° and 300° C. shows faint points of inflexion, corresponding respectively to metallic nickel contents of about 5 and 35 per cent.; the curves for reduction at 350° and 400° C. are quite smooth up to more than 90 per cent. free nickel.

With the fused oxide, a change of curvature is noticeable at about 80 to 85 per cent. reduction at a temperature of 500° C.

Nickel hydroxide on kieselguhr reduced at 500° C. gives a smooth continuous curve, inclining more rapidly to the abscissa axis after 50 per cent. reduction has been reached; at about 60 per cent. reduction, also, the activity of the product has attained a maximum value which is thereafter constant (*vide infra*).

The reduction curves are evidently related to the physical conditions (state of the oxide and temperature of reduction) rather than to the formation of any more or less definite sub-oxides intermediate between nickelous oxide and nickel.

Relation of the Activity of a Supported Nickel Catalyst to its Reduced Metal Content.

The extent to which the solid surface most easily accessible to gases or liquids determines the extent of the activity displayed is well illustrated by overloading a support with a nickel compound and observing the activity of the products formed by gradual reduction of the latter.

Portions of a preparation of nickel oxide on kieselguhr containing a total amount of 14·62 per cent. nickel (calculated on the dry fully reduced material) were reduced in hydrogen at 350° to 500° C. to varying extents and the relative activities of the products measured roughly by the melting points of the fats produced when cotton seed oil was hydrogenated at 180° C. for one hour with an amount of the catalyst such that the total nickel present was 0·1 per cent. of the oil. The following results were obtained:—

Percentage reduced nickel.	Ratio of reduced : total nickel.	M.P. of hardened fat.
		° C.
1·89	0·129	37
3·85	0·263	50
5·57	0·381	54
7·43	0·508	59
9·16	0·631	59½
11·80	0·807	60
13·01	0·890	59½
14·21	0·972	59½

In another series of experiments the catalytic activities were compared in terms of the absorption of hydrogen per minute during the first phase of hydrogenation of the glycerides of linseed oil: we have shown* that in this stage of the hydrogenation process the graph relating hydrogen absorption to time is almost entirely linear.

Two hundred grm. of highly refined linseed oil were hydrogenated at 180° C. in presence of 2 grm. of catalyst containing a total amount of 14·50 per cent. nickel (calculated on the dry reduced catalyst), the total concentration of nickel present in the case of the fully reduced catalyst being therefore 0·145 per cent. of the weight of glyceride:—

Percentage reduced nickel.	Ratio of reduced : total nickel.	Catalytic activity.	
		H_2 absorption litres per minute.	Ratio (most active catalyst as unity).
		Measured over 11 litres.	
2·88	0·199	0·130	0·280
3·56	0·246	0·159	0·342
5·63	0·389	0·302	0·649
6·45	0·445	0·333	0·716
8·46	0·583	0·456	0·981
10·36	0·714	0·465	1·000
12·85	0·886	0·465	1·000
14·19	0·979	0·445	0·957

It is considered that figures of this nature indicate that the activity is dominated by the condition of the surface layer of reduced nickel, since it is obviously this part of the mass which will be first reduced by hydrogen to the metallic state.

It is notable that, as in the case of "non-supported" nickel catalysts, maximum activity is reached when only a portion of the nickel oxide has been reduced to the metallic state, but in the present case the support itself is not appreciably affected in bulk by the process of reduction, and, instead of declining again, the catalytic activity is maintained constant until and when the whole of the nickel oxide present has been reduced.

Comparison of the activity of the supported and non-supported catalysts, in the light of the parallel observations on the bulk gravities of the materials, leads inevitably to the conclusion that these variations in catalytic power are to be ascribed simply to the different surface areas of free nickel exposed, and do not require for their interpretation the assumption of the presence of any catalyst other than metallic nickel.

---

* 'Roy. Soc. Proc.,' A, vol. 96, pp. 137–146 (1919).

Anomalous Liesegang Stratifications produced by the Action of Light.

By EMIL HATSCHEK.

(Communicated by Prof. A. W. Porter, F.R.S. Received May 27, 1921.)

(PLATE 8.)

The investigation to be described in the present paper was originally begun with the intention of studying the Liesegang phenomenon under conditions more strictly defined than have been observed hitherto, and more particularly with definite volume ratios of the reacting solutions. In all the experiments recorded in the literature the amount of solution placed on a given, or unknown, quantity of gel has been arbitrary and is generally not even stated. A quantitative treatment of the phenomenon is impossible *a priori* as long as this ratio is unknown, and would probably become simplest if one, or both, volumes were made—practically—infinite. In the course of the preliminary studies for such experiments an extremely striking anomaly, not observed so far, was discovered, which appears to deserve being put on record before the main investigation is completed.

The reaction chosen for this work was the formation of lead chromate in agar gel by the interaction of lead acetate (dissolved in the gel) with potassium or sodium dichromate, or with potassium chromate. This reaction, in very low concentrations, produces extremely fine and numerous stratifications, which were first obtained by L. J. de Whalley;* the low concentrations and the cheapness of the ingredients also made it eminently suitable for work involving the use of large volumes. Soon after its discovery the reaction was employed by the author and the results discussed in their bearing on some controversial points, in the paper quoted above. A normal specimen is illustrated in fig. 1 (Plate 8); in this, as in all other specimens photographed, the excess of soluble dichromate present in the gel has been removed by prolonged diffusion, so as to show the lead chromate strata as distinctly as possible.

No anomalies whatever were observed in any of the specimens obtained in 1914. When work on the present investigation was begun, the difficulty generally encountered in the study of the Liesegang phenomenon, viz., that of duplicating results with different brands of raw material, was experienced. This difficulty is generally much more acute with gelatin than with agar, but in the present case one brand of (shred) agar was found to give no strata at all, while another (powdered) gave results much inferior to those obtained in 1914. A very satisfactory sample of powdered agar was eventually found and a standard procedure of preparing the 1 per cent. gel used throughout was adopted and rigidly adhered to. The agar sol, cooled to about 90°, received an addition of 0·1 grm. of crystallised lead acetate ($PbAcet_2.3H_2O$) to every 100 c.c. of sol, and was poured into clean and carefully dried test-tubes. When the gel had cooled to room temperature, a solution containing 5 grm. per litre of crystallised potassium or sodium dichromate (0·5 per cent.) was carefully poured on the gel. As the sodium salt contains two molecules of water of crystallisation, while the potassium salt is anhydrous, the 0·5 per cent. solution of the former has a slightly lower equivalent concentration than that of the latter; nevertheless, it diffuses somewhat more rapidly into the gel than the potassium salt. Apart from this, the results are indistinguishable. The molar concentrations are thus: 0·1 per cent. of crystallised lead acetate = 0·246 millimoles, 0·5 per cent. of crystallised potassium dichromate = 1·70 millimoles. The component which is to diffuse into the gel is accordingly, as is quite generally the case, in considerable molar excess. It was later found that both concentrations could be doubled, and equally good results be obtained somewhat more rapidly. In some cases the gel was made acid by the addition of one to five drops of N/1 acetic acid to 10 c.c. of gel; it is not necessary to enter into the effect of this addition beyond saying that the strata become somewhat more substantial and slightly further apart.

---

* Hatschek, 'Koll.-Zeitschr.,' vol. 14, p. 115 (1914).

In the initial experiments test-tubes containing generally 10 c.c. of gel were used, and measured volumes, varying from 1 to 4 c.c., of dichromate solution were placed on top of the gel. The test-tubes were placed in the ordinary racks on the laboratory table. With this arrangement, normal stratifications were obtained only for a short distance in the top portion of the gel column; as diffusion progressed, a notable change occurred, wide bands of precipitate being formed at intervals between the normal strata (fig. 4). A striking feature of these anomalous bands is that the distance between adjoining ones *decreases* as the reaction proceeds, while the distance between the normal strata always increases. Repetition of the experiment at intervals led to substantially the same result, although the number of normal strata in later specimens decreased noticeably (fig. 5). A very curious feature was observed in the sample of agar, mentioned above, which failed to produce normal strata at all; the abnormal bands are nevertheless well marked in the uniform precipitate (fig. 6).

The first explanation of the anomaly which suggested itself was the rapid decrease in the concentration of the incoming reaction component, with the particular ratios employed, although it must be mentioned at once that no such abnormal formations have ever been described in the case of other reactions carried out with quite arbitrary volumes of solution and gel. On the other hand it had to be considered that the concentrations used were exceptionally low compared with those employed in most other reactions; thus the lead iodide reaction gives the best results in 1 per cent. agar gel containing 2 to 4 per cent. of KI, covered with a 30 per cent. solution of $Pb(NO_3)_2$, *i.e.*, with molar concentrations over 500 times as great as those used in the lead chromate reaction.

To eliminate this possibility it seemed advisable to return to the original object of the investigation and to use a volume of dichromate solution so large in comparison with that of the gel that the decrease in concentration caused by the diffusion into the latter and by the reaction would be negligible. The test tubes containing the agar gel were accordingly immersed completely in dichromate solution, stirred at intervals and contained in glass jars from 300 to 1,500 c.c. capacity. All these produced very perfect stratifications of the normal type: in those containing the slight addition of acetic acid mentioned above the curious "faults" and anastomoses between adjoining strata, which are highly characteristic of agar* show themselves strikingly (fig. 2). In some specimens this joining of adjacent strata is so regular that the whole system forms a continuous pseudo-helical surface (fig. 3).

The strata show, from the top downward, the regular increase in distance which is generally, and no doubt correctly, ascribed to the gradually decreasing concentration of the component in the gel. It seemed desirable to proceed one step further and to use both a large volume of solution and a large volume of gel, in order to study the effect of keeping the concentrations of *both* components practically constant. The simple arrangement shown in fig. 7 was used for this purpose, with various unessential modifications. A tube A, about 1 cm. bore and filled with the lead acetate—agar gel dips into a flask containing from 120 to 200 c.c. of the same gel. The top of the tube carries a funnel, B, containing about 120 c.c. of dichromate solution. As the volume in a length of 20 cm. of tube is about 15 c.c., this arrangement keeps the concentrations of both lead acetate and dichromate constant within roughly 12 per cent.

Various experiments were carried out with this arrangement, all of which again produced the anomalous phenomenon, in an accentuated form; in the majority of cases the normal stratifications were confined to the top of the gel column (fig. 8). The assumption that the constant concentration of dichromate, in the jar experiments, prevented the anomaly, was therefore rendered untenable.

---

* *Cf.* E. Küster, 'Koll.-Zeitschr.,' vol. 18, p. 107 (1916).

At this stage it was decided to mark the progress of diffusion every twenty-four hours; the marks showing in fig. 8 were placed at 11 P.M. From this, and from numerous parallel specimens, it became quite evident that the abnormal strata were a *diurnal phenomenon*.

The only diurnal variation which suggested itself as likely to influence the phenomenon was that in temperature. In this connection it should be mentioned that the preparations were kept in two rooms with north aspect one of which was not heated while the other was warmed day and night. The temperature in the latter during the period covered by the experiments varied, roughly speaking, between 11° and 16° C., while that of the other was within 2° or 3° of the outside temperature. As each experiment extended over at least a week, and in many cases over two, the use of an efficiently stirred thermostat was inconvenient, and various forms of heat insulation were tried. The apparatus illustrated in fig. 7 was placed inside a tall cylinder filled with sawdust; in other cases only the tube A was surrounded by a wider tube, the space between the two being filled with sawdust. In other experiments test-tubes with only 2 to 4 c.c. of dichromate solution were placed in boxes filled with sawdust or shredded asbestos.

All these preparations, whether a small or large volume of dichromate solution was used, produced normal strata. Test-tubes immersed in a large volume of water, and fitted with funnels containing various volumes of dichromate solution, however, persistently produced anomalous results, although the heat insulation in this arrangement should have been equally effective as when the tubes were submerged in the dichromate solution itself.

The only possible conclusion was that the anomaly was produced by light. Dichromate solution, even at a concentration of 1 per cent. and in very moderate thickness, cuts off the rays of shorter wave-length than about 5,200 very completely, so that the test-tubes submerged in it were exposed only to the non-actinic portion of the spectrum. In all the other forms of heat insulation used light was also incidentally excluded.

This unexpected conclusion was verified by carrying out reactions in pairs of preparations, one of which was exposed to diffuse light, while the other was protected by a wrapper of thin copper or tin-foil, these materials being chosen so that variations in temperature should affect both preparations equally. The results leave no doubt regarding the effect of light. The screened specimens developed normal strata throughout, whereas those exposed to light all showed anomalies. The character of the latter varied considerably, a feature which was to be expected, since the experiments extended from March to May, and the hours of daylight as well as the intensity of the latter varied considerably over the period. In all cases, however, the *formation of normal strata was interrupted* when illumination was sufficient, and their place was taken by wider bands which sometimes differed markedly in colour from the normal strata and were often perceptibly denser on the side which had been turned towards the window. In the specimens obtained later in the year normal stratification generally stopped altogether after the first forty-eight hours, and only wide diurnal bands were formed after that.

It is highly interesting that both the normal and the abnormal bands can be produced side by side in the same body of gel: two specimens showing this result are illustrated in figs. 9 and 10. The gel in both cases was contained in a cell only 4 mm. deep; one half (the left one) of the face turned towards the light was covered with black paper, while the back was completely covered, so as to prevent reflection. In both cases, the anomalous strata in the right half—that exposed to light—show very clearly. The normal strata show only slightly in fig. 9, this preparation having been made with a very small volume of dichromate solution, whereas they are clearly visible throughout the left half of fig. 10, made with a large volume of solution. The irregular appearance is generally to be observed in preparations made in very shallow cells, and may be partly due to the meniscus and partly to inequalities in the agar due to rapid cooling. The progress of diffusion in fig. 10 was marked during the first three days at 9 A.M. (full lines) and 9 P.M. (dotted lines). It is worth mentioning that this preparation was made about the middle of May, when the window, which looks almost due north, received some direct light before 7 A.M. and after 7 P.M. (summer time).

At the present stage it is impossible to suggest any explanation of this very striking effect of light, but a few possibilities, which rather readily obtrude themselves, may be ruled out. The occurrence of the phenomenon in an organic medium containing a dichromate seems to indicate a clue, but the attempt to trace the anomaly to a possible reducing action of the organic gel under the influence of light does not lead very far towards a satisfactory explanation. Such an action, though it has never been recorded with agar, is not *primâ facie* improbable, as it occurs in other carbohydrates, *e.g.*, gum arabic. To test this point, 1 per cent. dichromate solution was allowed to diffuse into duplicate specimens of pure agar gel, *i.e.*, without any lead salts, one specimen being exposed to light and the other covered with black paper. No difference in the rate of diffusion could be detected, nor in the appearance of the specimens: the colour in both was that of pure dichromate. Both specimens, on being slowly heated together in the water bath, "melted" at the same temperature and were equally clear.

No attempt was made to separate any possible insoluble product of reduction by filtration or sedimentation, as either would have involved keeping the agar sol at about 90° C. for some time, a procedure likely to lead to reduction by itself.

It must also be remembered that the salt present in the gel throughout is *not* the dichromate, but the lead acetate, which appears extremely unlikely to be acted on by light. Gels containing lead acetate have been repeatedly kept, fully exposed to light, for several days, have been re-melted, and have eventually produced quite normal strata in test-tubes protected from light. The dichromate only diffuses in gradually, and it is very difficult to imagine that any reaction between it and the agar, under the influence of light, could proceed nearly as rapidly as the ionic reaction between dichromate and lead acetate; in other words, the precipitate of lead chromate should be formed long before any reduction of dichromate in the reaction zone, assuming it to take place at all, would seem probable. Of course, these considerations do not exclude the possibility that some product is formed which inhibits normal stratification, and that this product diffuses in advance of the dichromate. In view of the known extreme sensitiveness of the phenomenon to exiguous amounts of admixtures, this possibility must not be neglected, although it is not capable of exact formulation at the moment.

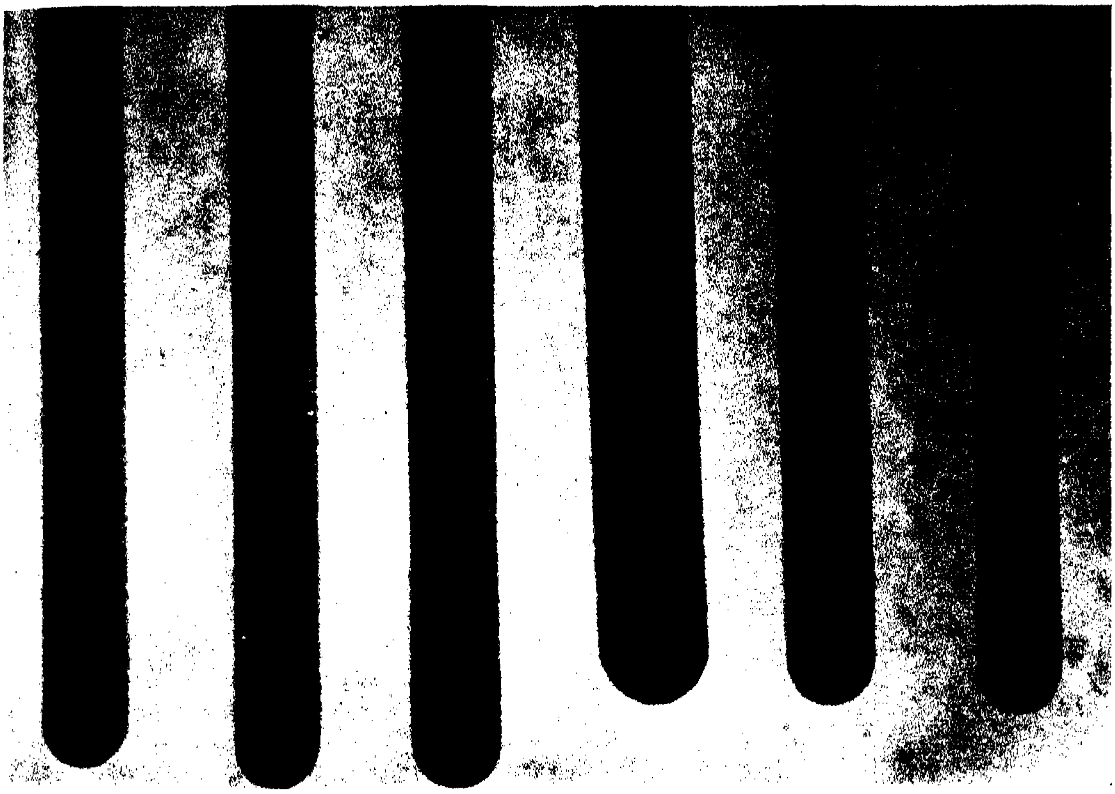
In this connection it is of interest to observe, more especially in fig. 10, that some traces of anomalous formation can be seen in the screened (left) half of the preparation. If such an inhibiting product were formed, it would of course diffuse in all directions, and would, therefore, also reach the portion of the gel protected from light. While this cannot be overlooked, an explanation decidedly more probable seems to be that scattered or reflected light has reached the shaded portion; although the back was covered with black paper, this was naturally on the outside of the glass wall, so that reflection from the internal face was still possible.

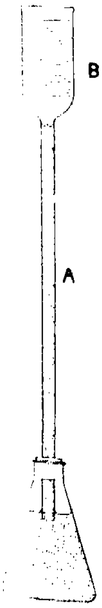
Finally, it must be mentioned that both the normal and the anomalous phenomenon can be obtained *equally well with chromate* as with dichromate. No data on the photo-chemical behaviour of chromate-carbohydrate mixtures seem to be available, but it is known that chromate-gelatin, unlike dichromate-gelatin, is *not* sensitive to light.

In the absence of other clues it seemed possible that the size of the grains of precipitate might throw some light on the mechanism of the anomaly. Several specimens, duplicates of those illustrated, were accordingly removed from the test-tubes—which can be done conveniently, as agar does not adhere to glass—and the precipitate from various regions examined and measured with the eye-piece micrometer. The precipitate, whatever its size, consists of roughly spherical granules, pairs of them being, very occasionally, joined to form a squat "hourglass" aggregate. As regards the size, the results are not easy to interpret. They are simplest in the specimen shown in fig. 6, which did not produce normal strata at all, but only the anomalous bands caused by light embedded in a uniform precipitate. The latter consisted of granules of remarkably uniform size, 5μ average diameter, while the bands contained equally uniform granules about $2 \cdot 5\mu$ average diameter. A duplicate of fig. 5 was also investigated, samples of precipitate being taken from three portions: A, from the first set of normal strata at the top of the gel, B from the first wide daylight band, and C from the three normal strata below this. A consisted of granules 1·25 to $1 \cdot 8\mu$ dia., B about $2 \cdot 5\mu$ dia., and C 3·7 to 5μ dia.

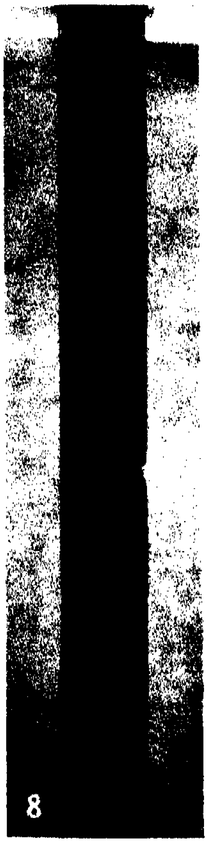
The latter results are difficult to interpret definitely; as regards the specimen, fig. 6, it is evident that the number of centres of crystallisation in the daylight bands is considerably greater than in the precipitate formed in the dark, or that the limit of supersaturation is altered. The question what causes this difference of course remains unanswered: a direct effect of light, in the absence of parallel cases, would have to be proved by more substantial evidence, while the formation of some reaction product which affects supersaturation or the number of nuclei is possible, but equally in need of proof. Further research must be directed to these two points, and the first step seems to be the elimination of the organic matter, by studying the formation of lead chromate, either in an inorganic gel, *e.g.*, silicic acid, or possible merely in aqueous solutions. It will also be necessary to ascertain whether the effect is peculiar to this one reaction, or whether it is of a more general nature.

In the meantime it is permissible to draw attention to one point, even in the absence of any explanation of the effect. The possible bearing of the Liesegang phenomenon on the formation of rhythmic structures in nature, and in conditions where there is no obvious rhythm or periodicity in the supply of reacting material, has been repeatedly pointed out, by Liesegang himself, by H. Bechhold, D'Arcy W. Thompson, E. Küster and others. If the diurnal change in illumination is capable of modifying profoundly the stratifications produced by a given reaction, it is evident that structures of much greater complication than those contemplated by the authors quoted may conceivably be the result of simple and non-periodic diffusion.

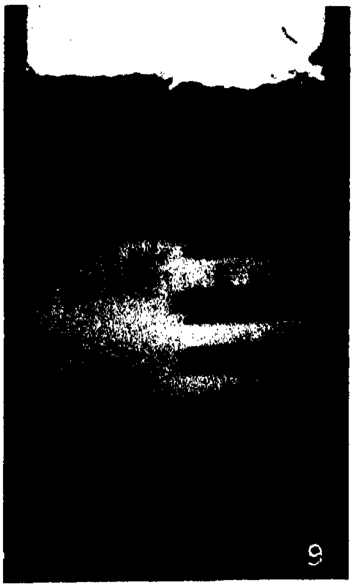




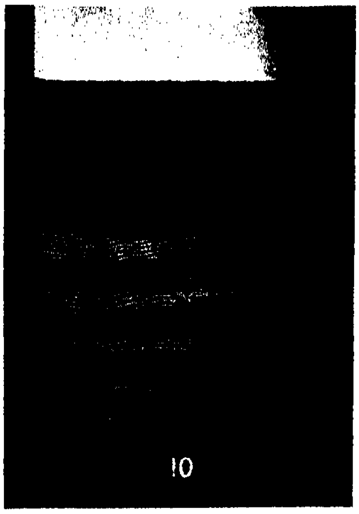
7.



8



9



10

On the Diffraction of Light incident at nearly the Critical Angle on the Boundary between two Media.

By Brojendra Nath Chakravarty, M.Sc., Assistant to the Palit Professor of Physics, University of Calcutta.

Communicated by Prof. J. W. Nicholson, F.R.S. Received January 12, 1920,—Revised February 20, 1921.)

[Plate 9.]

1. *Introduction.*

When a parallel pencil of monochromatic light falls at nearly the critical angle on the plane boundary between two media of differing refractive index, and emerges in the rarer medium in a direction almost parallel to the surface of separation, the aperture of the emergent pencil becomes greatly restricted. A study of the diffraction phenomena observed in this limiting case of oblique emergence seems of interest, especially in view of the fact that in the Lummer-Gehrcke interference spectroscope we have an actual case of importance in which such obliquely emergent pencils determined the observed phenomena. Some preliminary work on this subject was carried out by Prof. C. V. Raman, and at his suggestion the author of the present paper undertook a detailed experimental investigation of the case with a view to establishing its principal features. So far as the writer is aware, though the theory of the parallel-plate interferometer in its essentials is well known,* the special character of the diffraction phenomena observed with pencils very obliquely refracted at a plane surface had not been discussed previously. It is proposed in the present paper to describe the results obtained by the writer, some of the observations (to be mentioned below) having special reference to the theory of the Lummer-Gehrcke plate.

2. *Experimental Methods and Results.*

The general features of the case may be observed on an ordinary spectrometer. It is well known that an ordinary prismatic spectroscope may be so adjusted as to give very large dispersions, by putting the prism on the table of the instrument in such a position that the light incident on the first face, after passing through it, falls at nearly the critical angle of incidence on the second face and emerges nearly grazing the surface. The lines in the spectrum then appear very widely separated, and also strongly curved, owing to the fact that pencils of light from different points on the slit of the collimator do not all pass through the principal plane of the prism. In his investigation of the optical power of spectroscopes, Lord Rayleigh* has remarked that the resolving power of the prism in the position referred to above is no greater than at other positions. The width of the beam emerging from the prism tends to zero as the grazing position is approached, and the spectrum lines are much widened by diffraction. This sets off the effect of the increased dispersion on the resolving power. Owing to the great dispersion and the enfeeblement of the light in oblique transmission, a very powerful source of monochromatic light is necessary for satisfactory observation of the diffraction phenomena. A Westinghouse quartz silica lamp is most suitable, though visual observations may also easily be made with a glass Cooper-Hewitt mercury-vapour lamp. Observations are made on the green line of the spectrum. In order to secure an extended diffraction-pattern, it is advantageous to use something less than the full aperture of the prism to start with, cutting it down by one or more slits having their edges parallel to the slit of the collimator. The position of the slits is a matter of indifference, so long as the aperture of the beam is cut down before it emerges or just as it emerges from the second surface of the prism, and not *after* it has emerged. The simplest arrangement is to limit the aperture of the beam by a slit or slits immediately before entry into the prism. Another arrangement would be to use a hollow prism, filled with liquid, and to put the slit or slits inside the prism, so as to limit the aperture of the beam just before incidence on the second face. A third arrangement is to use a glass prism the aperture of the second face of which has been reduced by grinding down to the desired width with parallel edges. All the arrangements give identical results.

---

* Lummer and Gehrcke, 'Ann. d. Physik,' vol. 10, p. 457 (1903); also 'Phys. Techn. Reichsanstalt, Wiss. Abh.,' vol. 4, pp. 63–84 (1904); Kolacek, 'Ann. d. Physik,' vol. 39 pp. 997–1041 (1912); Gehrcke, 'Anwendung der Interferenzen,' Braunschweig, 1916.

Several photographs of the diffraction-pattern seen in the observing telescope were secured, of which a selection is reproduced as figs. 1, 2, and 3 in the plate. Fig. 1 represents the effect observed with a single slit limiting the aperture of the beam, the prism being in a position very nearly that of grazing emergence. It will be noticed that the diffraction-pattern is quite unlike the symmetrical system of bands usually obtained on the spectrometer with a single slit. Apart from the curvature of the bands (which has already been explained above), the pattern is markedly asymmetrical in character. To the left of the central bright fringe we have a regular succession of bands slowly decreasing in width and brightness, while on the right we have only one band, which is broad and very faint compared with the first band on the left. In general, it is found that the fringes to the right are fewer in number, broader, and much fainter than the corresponding fringes on the left. With the prism very near the critical position, the fringes to the right may even entirely vanish.

---

* Rayleigh, 'Scientific Papers,' vol. 1, p. 426.

Figs. 2 and 3 represent the effects observed when the aperture of the prism is restricted by a row of equidistant parallel slits, forming a grating, the prism being nearer the critical position for fig. 3 than for fig. 2. It will be noticed that the asymmetry is very marked in these cases as well, the difference in the intensity of diffracted images of the slit to the right and to the left of the central group being specially evident in fig. 3.

3. *Photometric Study of the Diffraction Pattern.*

To put the foregoing observations on a quantitative basis, a set of photometric comparisons of the intensities of the corresponding bands of the same order on the two sides of the pattern was carried out. For this purpose, a rotating sector photometer, of a modified Abney type, was used. It was supplied by Messrs. Adam Hilger, the special feature in it being that the free disc (which can be adjusted by a handle while in rotation) is smaller in radius than the fixed disc. The sectors, while in rotation, thus present two annuli of different intensities, the ratio of which can be adjusted at pleasure by moving the handle of the instrument. The disc of the photometer is placed at the focal plane of the objective of the observing telescope, so that the diffraction pattern can be seen through it with an eyepiece, the fringes on the brighter side being observed through the inner annulus of the disc, and those on the fainter side through the outer annulus. To enable the intensities at corresponding points on the two sides of the pattern to be compared, a screen with two vertical slits is interposed immediately in front of the photometric disc, so as to cut off everything except the regions under observation, which are then adjusted to equality of brightness by moving the handle of the rotating apparatus. Several readings can be taken in succession and their average found. The angles of diffraction, θ and θ' for the two bands under comparison, can be found by reading their positions, and the position of the second surface of the prism on the divided circle of the spectrometer.

The results obtained with a single slit, limiting the aperture of the beam, are shown in column 3 of Table I. Column 1 shows the serial number of the experiment, the position of the prism on the table of the spectrometer in the successive experiments being altered so as gradually to decrease the obliquity of the emergent pencil of light.

Table I.

1.	2.	3.	4.	5.
Serial No. of experiment.	Serial order of the maxima compared.	Observed ratio of intensities.	$\cos^2\theta/\cos^2\theta'$.	No. of maxima visible to the right.
1	First	4·74	4·95	1
2	First	2·55	2·52	2
2	Second	4·74	5·15	2
3	First	1·96	1·65	3
3	Second	2·56	2·54	3
3	Third	3·46	3·87	3
4	First	1·84	1·61	3
4	Second	2·30	2·18	3
4	Third	3·10	3·20	3
5	First	1·41	1·36	3
5	Second	1·84	1·74	3
5	Third	2·43	2·24	3

A few determinations were also made with a wire-grating, limiting the aperture of the beam entering the prism. The results are shown in Table II.

Table II.

1.	2.	3.	4.
Serial No. of experiment.	Serial order of maxima compared.	Observed ratio of intensities.	$\cos^2\theta/\cos^2\theta'$.
1	First	5·29	5·55
2	First	3·21	3·12
3	First	2·81	3·02

In these cases, the distances between the diffracted images of the higher orders on the two sides of the pattern were too great to enable them to be seen simultaneously in the eyepiece. Consequently, the photometric observations were not extended to higher orders than the first.

4. *Apparent Failure of Snell's Law near Grazing Emergence.*

An interesting consequence of the markedly asymmetrical distribution of intensity in the diffraction pattern is that the position of maximum intensity in it is not given by the ordinary law of refraction, but is appreciably displaced to one side, so that the angle of diffraction for this position is less than the angle of refraction as given by Snell's law. The magnitude of the displacement is greatest when the emergent light practically grazes the surface of the prism, and also depends on the aperture of the pencil, being greatest when the aperture is smallest, and *vice versâ*. The displacement can be measured by first observing the position of maximum intensity when the full aperture of the prism is used, and then also when the aperture is reduced to any desired extent by the interposition of a slit before the first face of the prism. The displacement of the position of maximum intensity being negligible for the full aperture, the difference of the readings gives the magnitude of the displacement when the smaller aperture is used. On account of the great width of the central band of the diffraction pattern in such cases, it is not easy to fix very precisely upon the position of maximum intensity by direct eye-observation. Nevertheless, the displacement is clearly detectable, and can be roughly measured.

Table III shows for various apertures, the angle of refraction as given by Snell's law, and (in column 3) the position of maximum intensity in the pattern as actually observed.

Table III.

1.	2.	3.	4.
Aperture.	Angle of refraction by Snell's law.	Angle of diffraction for the central maximum. (Observed.)	Angle of diffraction for the central maximum. (Calculated.)*
mm.	° ′	° ′	° ′
1·90	88 39	88 37·2	88 37
0·70	88 20·6	88 17·8	88 17
0·60	88 14	88 10·8	88 10
0·42	88 35·4	88 31·1	88 30·6
0·30	88 25	88 19·4	88 16·8

* See Section 6.

5. *Diffraction of Light in the Lummer-Gehrcke Plate.*

As is well known, in the special form of interferometer developed by Lummer and Gehrcke, light enters a plane-parallel plate, and suffers a series of internal reflections at nearly the critical angle of incidence, and, at each such incidence, part of the light emerges as a thin pencil in a direction nearly grazing the surface of the plate. It is obvious that, if only one of such pencils entered the observing telescope, we would see in the field of view an unsymmetrical diffraction pattern of the type described in the preceding sections. Actually, the effects seen would be modified by two causes: firstly, by the interference (under large difference of path) of the several pencils emerging from the plate; and, secondly, when an extended source of light is used, by the superposition of the effects of the different parts of the source. The second complication may be avoided by narrowing down the slit of the collimator from which the light enters the plate, and setting it parallel to the plane of the plate. The interference maxima seen in the field of view would then be formed entirely by the light diffracted on emergence from the plate. The unsymmetrical character of the distribution of intensity, and its similarity with the phenomena previously described, is rendered most conspicuous by somewhat reducing the aperture of the pencil immediately before its entry into the plate, so that the number of interference maxima falling between consecutive diffraction minima becomes fairly large.

Figs. 4, 5 and 6 show the effects then observed for different angles of incidence, and make evident the great difference in intensity of the corresponding interference maxima on the two sides of the central fringe. The difference can be seen even when the full aperture of the instrument is used. This is clear from figs. 7 and 8 in the Plate, which also serve to illustrate another important feature in the theory of the Lummer-Gehrcke plate. It will be noticed that each of these figures shows the interference maxima for two different angles of incidence of the light, and the positions of these maxima, formed in diffracted light, remain unaltered—though their intensities change—when the angle of incidence is varied. It is this invariability in the positions of the interference maxima that enables an extended source of light to be used in practical work with the Lummer-Gehrcke plate without resulting in loss of sharpness of the fringes. It also explains why in figs. 4, 5 and 6 the interference fringes run practically straight, while the diffraction maxima and minima are strongly curved.*

6. *Explanation of the Results.*

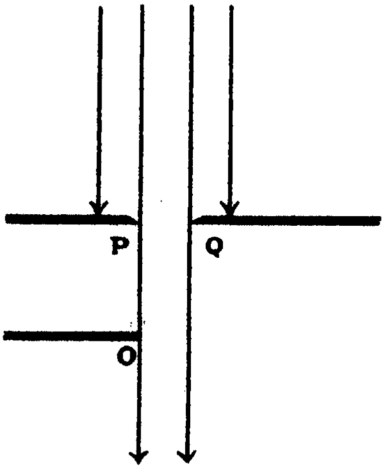
(*Rewritten February* 20, 1921.)

It may be remarked that diffraction-phenomena analogous to those described above are also noticed when the pencil of light from the collimator is reflected at nearly grazing incidence from the surface of a prism and is viewed through the observing telescope. Indeed, similar effects may also be obtained very simply by merely placing two razor-edges facing opposite ways, one nearly behind the other, parallel to the spectroscope slit, on the table of the instrument and viewing the light transmitted through the very oblique aperture thus formed. In the diffraction pattern obtained by any of these methods, the fringes on the left are fewer, broader and much less intense than the fringes on the right of the central maximum, in the same way as in the case at present under consideration.*

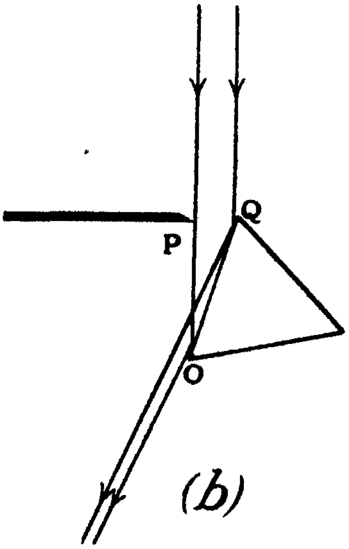
---

* That the positions of the interference maxima formed by the light diffracted on emergence from the plate remain unaltered when the incidence of the light at entry into the plate is varied, may be explained in a very simple manner. From the principle of minimum or stationary path (Fermat's Law), it follows that the retardation of a wave which starts at a given position and arrives at a specified point in the focal plane of the observing telescope, after one or more reflections within the plate, depends only on the number of such reflections, and is independent of any variations in its actual path, so long as such variations are small. It follows that the positions of the interference maxima are unaffected by small changes in the actual path of the light.

The referee of the Royal Society in commenting on the explanation of the phenomena put forward in this paper, has suggested two experiments which have also been tried by the author. These experiments are indicated diagrammatically in figs. (*a*) and (*b*). In the first, three razor-blades P, Q, O, were placed on the table of the spectrometer as shown, and the light transmitted through the arrangement was viewed in focus through the observing telescope.



(*a*)



(*b*)

The second experiment, fig. (*b*), consists in allowing a pencil from an aperture limited by a diaphragm to fall obliquely on a reflecting surface OQ. In both experiments, however, somewhat complex effects were noticed, which varied with the distances between the three edges or with the position of the diaphragm and the mirror and which perhaps can be best described as consisting of two or more diffraction-patterns super-imposed on each other. This is not surprising, as in figs. (*a*) or (*b*) the third edge or diaphragm introduces an additional complication which modifies the effects due to the obliquely held aperture or reflecting surface taken by itself. In any case, however, these observations did not prove to be of assistance in clearing up the nature of the effects under discussion noticed by the author.

---

* The effects observed with oblique reflexion and oblique transmission have been studied by C. V. Raman, 'Phil. Mag.,' November, 1906, January, 1909, and May, 1911. See also S. K. Mitra, 'Phil. Mag.,' January, 1918.

In all the cases with which we are now concerned, the common feature is that the fully illuminated aperture from which the light finally emerges, the elements of which should be regarded as the centres of diffracted disturbance, lies in a direction nearly parallel to the direction of the diffracted rays. Dividing this aperture into equal small elements and considering the phases of the beams incident on and diffracted from these elements and summing their effects, we find that in directions given by the formula

$$a(\mu \sin i - \sin \theta) = \pm n\lambda, \tag{1}$$

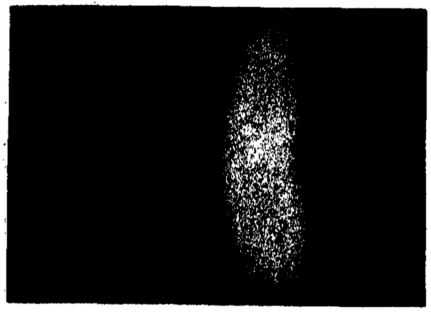
the illumination in the diffraction-pattern should be zero, i being the angle of incidence, θ the angle of diffraction, a the aperture and μ the refractive index. The intensity at any point in the diffraction-pattern as obtained by summation of the effects of the elements of the aperture is given by

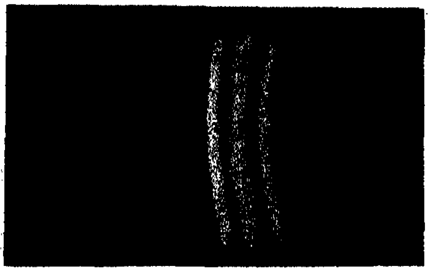
$$\mathrm{I} = \mathrm{I}_0 \sin^2 \phi / \phi^2, \tag{2}$$

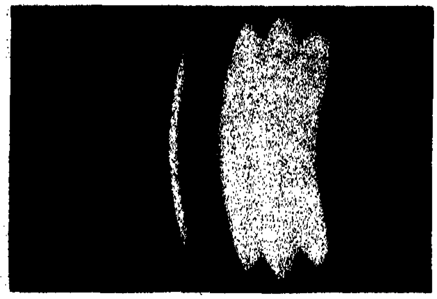
where $$\phi = \frac{\pi a}{\lambda}(\mu \sin i - \sin \theta).$$

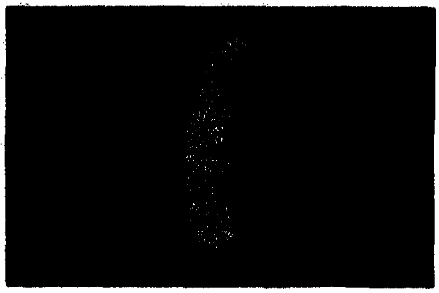
In agreement with these formulæ, we find that the dark bands in the diffraction-pattern are sensibly of zero illumination and their positions agree with theory. In the cases under consideration θ is nearly equal to $\pi/2$, and it is readily shown that the angular separation of the dark bands should increase continually from one side of the pattern to the other, exactly as observed. It is also clear why there are fewer bands on one side, as the value $\theta = \pi/2$ sets a limit to the extension of the pattern in that direction.

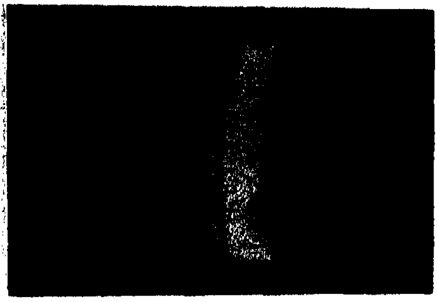
The feature requiring special explanation is the asymmetry in the *intensity* of the bright bands on the two sides of the pattern. If in formula (2) the quantity I_0 were a constant with respect to θ, the distribution of intensity in the pattern would be symmetrical, that is, corresponding bands on either side of it would be of equal intensity. Actually, however, this is not the case, as we have seen, and we are led to regard the observed effects as indicating a variation of the quantity I_0, with the angle of diffraction θ. Such a variation might have been expected, *a priori*, as the direction of the diffracted rays is nearly parallel to the aperture, and hence the cross-section $a \cos \theta$ of the diffracted beam varies rapidly with θ (which is nearly equal to $\pi/2$), instead of being practically constant as in the case of diffraction by a normally held aperture. The quantity I_0 is in fact proportional to the square of the amplitude of the secondary disturbance contributed by each element of the

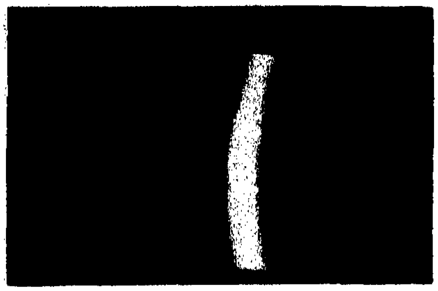

1

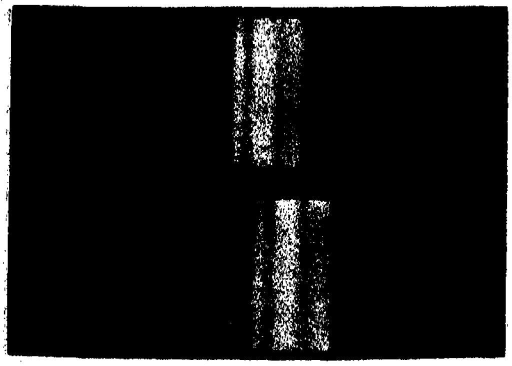

2

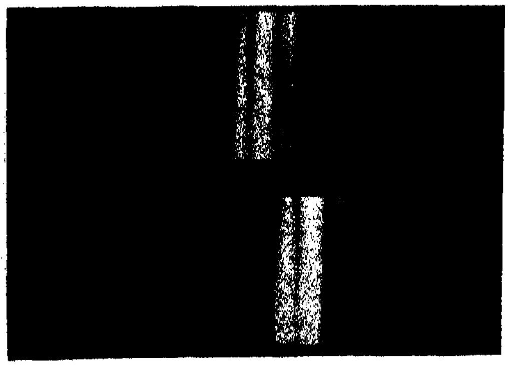

3


4


5


6


7


8

aperture, and if we assume this amplitude to be proportional to the cosine of the angle of diffraction θ, we obtain finally

$$I = I_0 \sin^2\phi/\phi^2 = I_1 \cos^2\theta \sin^2\phi/\phi^2,$$

where I_1 is sensibly a constant within the diffraction-pattern. On this basis, the ratio of the intensity of corresponding bands on either side of the pattern should be $\cos^2\theta_1/\cos^2\theta_2$ where θ_1 and θ_2 are the respective angles of diffraction, and this, as we have seen agrees with the results of the photometric determination (Section 3, above).

The foregoing also enables us to calculate the displacement of the position of maximum intensity in the pattern from that given by Snell's law. We may obtain the position of maximum intensity from the expression for I as given above, by putting $dI/d\theta$ equal to zero. It is found that for this position, the phase-angle ϕ is not zero, but is approximately given by

$$\phi = \frac{\lambda}{\pi a} \cdot \frac{3 \sin\theta}{\cos^2\theta}.$$

The angles of diffraction for the position of maximum intensity in the pattern have thus been calculated and shown in column 4 of Table III for comparison with experiment.

In conclusion, the author wishes to express his cordial thanks to Prof. C. V. Raman, who suggested the investigation and throughout took much interest in its progress, besides providing the necessary facilities for carrying it out at the Palit Laboratory of Physics.

OBITUARY NOTICES

OF

FELLOWS DECEASED.

CONTENTS.

PAGE

SIR W. DE W. ABNEY (with portrait) i

ROBERT BELLAMY CLIFTON vi

SIR LAZARUS FLETCHER (with portrait) ix

SRINIVASA RAMANUJAN xiii

WOLDEMAR VOIGT xxix

SIR W. DE W. ABNEY, K.C.B., 1843—1920.

By the death, on December 2nd, 1920, of Sir William Abney there passed away one of the notable figures of the scientific world of the past forty years and one who will be long remembered as the great pioneer in the science of photography.

That it was worthy to be ranked as a pure science he always insisted, and on more than one occasion he deprecated the attitude of scientific men who used photography as a mechanical aid in their laboratories without endeavouring to understand it, and lamented that of 25,000 people who took photographs not more than one cared for, or knew anything about, the why and wherefore. However this may be—and we must be careful not to push the contention too far, or we shall find ourselves committed to the proposition that a knowledge of the chemistry of ink is a necessary part of the equipment of a writer—there is no doubt that in Abney's hands photography was an exact science, an offshoot from both chemistry and physics, concerned with the action of light upon all manner of bodies. With this science and its growth over many most fruitful years Abney's name must always be inseparably connected.

He was the eldest son of Canon E. H. Abney and was born on July 24, 1843. He was educated at Rossall School and entered the Royal Engineers through the Royal Military Academy in 1861. The first few years of his service, spent partly in India, were uneventful. In 1870 he returned to England and was stationed at Chatham, and in the next year was appointed Assistant to the Instructor in Telegraphy at the School of Military Engineering. Here he found himself in charge of a small photographic establishment and chemical laboratory, then constituting part of the Electrical School, and at once began the active prosecution of photographic researches. He formed classes of officers and men for studying the subject, and the first edition of his book, 'Instruction in Photography,' destined afterwards to reach its eleventh edition and to be the guide of innumerable students of the art, was printed at Chatham in 1871 as a small pamphlet for the use of his pupils.

In the previous year he had joined the Photographic Society and his first paper, on "The Application of Albumen to Photography," appeared in 1870. Under his energetic direction the photographic establishment soon outgrew the tutelage of the Instructor in Telegraphy and in 1874 a separate Chemical and Photographic School was formed of which he was given sole charge. From this time onward he was so continuously engaged upon the evolution of photography into an exact science that it would be hard to find any side of this field of knowledge which he had not made a subject of experiment and upon which he had not written fully and critically.

In 1874 the only extensively used photographic process was the collodion "wet plate." The gelatine dry plate, though actually first made in 1871, was then very imperfect and the various dried collodion plates and collodion emulsions, in which Abney was one of the foremost experimenters, were only used by a few eager inquirers. In 1878-79, due mainly to advances made by Bennett and Abney in this country and van Monkhoven on the Continent, a rapid gelatine emulsion, in all main details identical with those used to-day, was first produced and the modern "instantaneous" photography made possible.

Abney rapidly established himself in a leading position as a practical exponent of the art and already in 1874 he was given complete charge of the arrangements for photographic observations of the Transit of Venus. He himself went to Egypt to observe it and brought back a great collection of views of temples and tombs, forming the basis of his book, 'Thebes and its Five Great Temples,' published in 1876.

Perhaps we may fairly claim that the first important memoir in which Abney showed his true quality and proved that he was indeed one of the elect to whom the why and wherefore are of the most basic importance, one who would never rest contented until he fully understood any process that went on under his observation, and saw that it is only by full understanding that further advance is made possible, was his paper "On the Alkaline Development of the Photographic Image" ('Phil. Mag.,' 1877). It had long been known—by whom the discovery was first made is still obscure—that where no free nitrate of silver was present, a considerably increased intensity of image was obtained with an alkaline developer, and it was generally assumed, though it cannot be claimed that there was much curiosity about it among the users, that the function of the alkali was to reduce the bromide, chloride, or iodide of silver, which had been acted upon by the light, to the metallic state. Abney showed, by a well planned and quite conclusive set of experiments, that the complete explanation was found in the fact that bromide of silver could not exist in close contiguity to the freshly reduced metal, but that sub-bromide was immediately formed, and similarly with chloride or iodide. This in its turn was again reduced to the metallic state by the developer, so that the image was eventually built up of metallic silver derived partly from those molecules of the salt acted upon by light, thereby reduced to the sub-salt, and further reduced to metal by the developer, and partly from the closely contiguous molecules, which were reduced to the sub-salt, not by light, but by contact with the freshly formed metal, and were then in their turn similarly acted on by the developer. The process was thus shown to be essentially different from the development of a plate with free nitrate of silver present, such as the ordinary wet-plate, where the salt reduced by light acts as a nucleus upon which metallic silver precipitated from the silver nitrate solution by the developer aggregates. As in the latter case the silver which forms the image is deposited out of a solution flowing freely over the plate there is almost no limit to the extent to which the image can be built up; the development and intensification can be pushed to the point where the high lights are completely opaque. Whereas, with an emulsion or dry-plate, the position of each molecule of silver salt is fixed, and it remains in the same place after reduction. If, therefore, the development is pushed too far, the image will spread out, losing form, and, should the action be allowed to continue to an extreme, the whole plate will ultimately become covered with an opaque mass of reduced silver.

In this power of building up the image without allowing it to spread lies the outstanding advantage of the wet-plate over the emulsion process and makes it, even now, when the manufacture of dry plates has reached such a high degree of technical excellence, preferable for some purposes.

In 1877 Abney left Chatham and joined the Science and Art Department at South Kensington, then under the able direction of the late Sir John Donnelly, R.E. In that department he served twenty-six years rising by successive steps to the post of Principal Assistant Secretary, Board of Education, which he held from 1899 to 1903, when he retired upon the reorganisation of the department under the late Sir R. Morant. He afterwards, up to the time of his death, held the honorary post of Scientific Adviser to the Board.

On taking up his duty at South Kensington he at once established a laboratory, placed in one of those hideous iron buildings known to two generations of Londoners as the "Brompton Boilers," and for nearly thirty years that laboratory was the source and fount of an unending stream of original experiment and research.

It would be impossible to follow Abney's work in any detail. His papers, as recorded in the 'Royal Society Catalogue,' number over one hundred, and to these must be added a very large aggregate of minor, but still important papers, over seventy in the 'Photographic Journal' alone, and many in the 'Journal of the Camera Club' and in other cognate publications. Much of the material of these memoirs and addresses was embodied by him in book form and his three standard volumes, 'Instructions in Photography,' 'Photography with Emulsions,' and 'A Treatise on Photography,' went through many editions.

One of his most noteworthy achievements was the photography of the infra-red region of the spectrum. He began experimenting upon this in 1875 while still at Chatham, and eventually succeeded in obtaining an emulsion of bromide of silver in collodion, in which the salt was in such a condition of molecular aggregation, indicated by its giving a blue colour to transmitted light, that it was sensitive to rays beyond the visible portion of the spectrum down to about λ 12,000. With this he mapped the solar spectrum from A to λ 10,650 ("Bakerian Lecture," 'Phil. Trans.,' 1880). In a subsequent paper ('Phil. Trans.,' 1881), written in collaboration with the late Maj.-Gen. Festing, R.E., he extended the use of these hypersensitive plates to the study of the absorption spectra of organic bodies in the infra-red, leading to important indications of their molecular groupings. For these researches he was awarded the Rumford Medal in 1882.

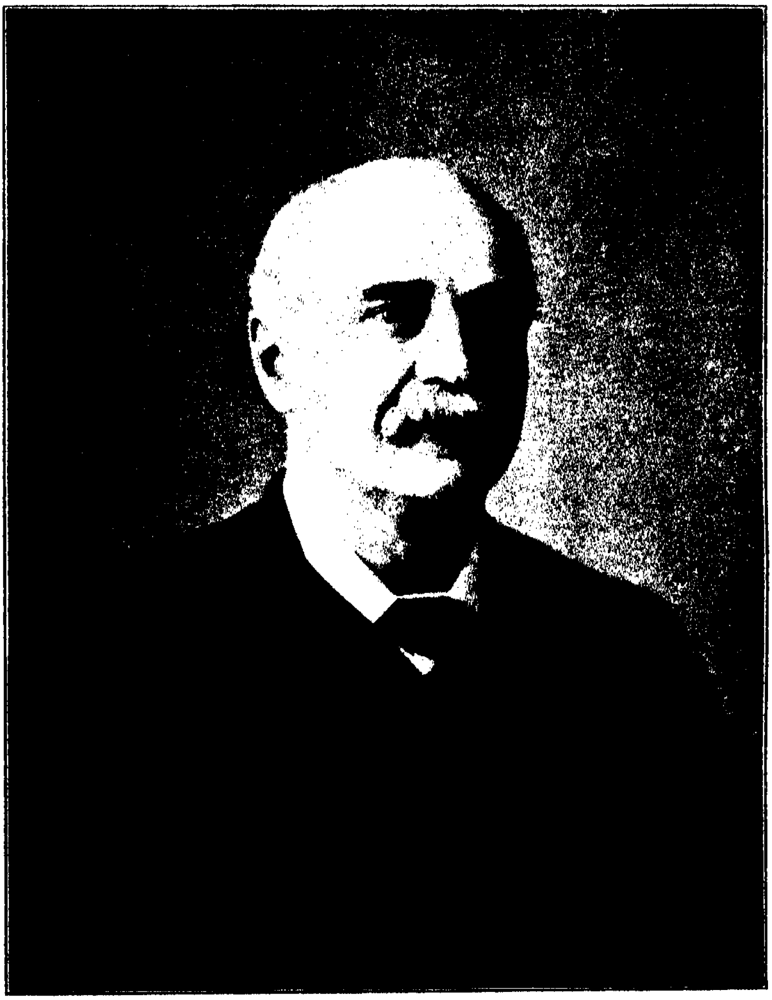
In this, as indeed in all his work, Abney showed a high level of manipulative skill. Many subsequent experimenters have tried to prepare an infra-red sensitive emulsion upon his formula and very few have succeeded. That it can be done there is however no doubt whatever; much depends upon the quality of the collodion.

His researches upon the action of the spectrum on the silver salts naturally led to the study of colour photography and in later years he did an enormous amount of exploration in this branch of the subject. In his earlier years he was very sceptical as to the possibility of any practical three-colour process; thus in 1878 he was of opinion that it was "utterly impossible to secure monochromatic colours which are pure enough to give the truths of nature," and that efforts (to obtain true colours by combining negatives taken through three different colour screens) "are not to be followed with too much zeal by scientific photographers." Later on, when it had been shown by Ives that exact colour representation could be obtained, provided the screens for taking the negatives and for viewing or projecting the positives were properly selected, he set out in a masterly way the principles involved in three-colour work. ("Theory of Colour Vision applied to Modern Colour Photography," 'Proc. Royal Institution,' 1899).

In the Bakerian Lecture of 1886 on Colour Photometry he gave an account of some experiments, made in collaboration with Gen. Festing, on the measurement of the relative illuminating intensities of the different parts of the spectrum. This subject had occupied him for some time previously and both then and afterwards, almost up to the time of his death, he devoted himself with immense enthusiasm to the problems of colour vision and colour blindness. He repeated, by a different method and in a far more exhaustive and complete form, the delineation of the intensity curves done by Clerk-Maxwell in 1855–6 and by König in 1883–1901. Without entering upon the vexed question of the true mechanism of colour vision it will suffice to state here that Abney's observations appear to confirm in every detail the intensity curves for the three fundamental colour sensations derived by König and thus afford the most powerful support to the validity of the Young-Helmholtz trichromatic theory of colour vision and the corresponding explanation of the different varieties of colour blindness. In addition to some thirty papers and memoirs he summarised his investigations and reviewed the whole subject in a book, 'Researches in Colour Vision and the Trichromatic Theory,' 1913.

Another notable research carried out partly in England and partly at the Riffel and other stations in the High Alps was embodied in his memoirs on "The Transmission of Sunlight through the Earth's Atmosphere" ('Phil. Trans.,' 1887 and 1893).

In 1882 he had planned to go to Egypt to observe the Total Solar Eclipse of May 17 but was prevented by temporary ill-health. This was a great disappointment to him. He was a keen traveller and invariably spent his summer in the Swiss or Italian Alps where he pursued his photographic work both from the scientific and the artistic side, and was indefatigable in water-colour sketching.



W. de W Abney

Abney was gifted in an unusual measure with true scientific insight, the power of seizing essentials, and we may almost say the power of predicting the future course of development. The point just mentioned of the three-colour process is probably the only one in which he was definitely wrong and where he failed to see the essentials of a problem. An extraordinary record for one whose output was so enormous! As an instance of his foresight it may be noted that in 1877, in a paper upon electric search-lights for military service, he advocated the employment of what is now known as a "flame" arc, using carbons charged with a calcium salt in order to give greater penetrative power to the beam by virtue of its ruddy colour.

All his experimental work was distinguished by its completeness and by the ingenuity and beauty of the means he devised. His early successes at Chatham in the improvement of photo-lithography; the invention of the "Abney ink"; the contrivance of the "Abney Level"; his methods for the photometry of both monochromatic and coloured lights; his precise measurements of the opacity of the deposit on the photographic plate leading to a statement of the law connecting density and exposure; his expedients for estimating the coronal light during a solar eclipse, for determining the transparency of the atmosphere and for the comparative evaluation of sunlight, starlight and skylight, were all marked with these same general features. He never used apparatus more elaborate than the actual conditions demanded; much of his work was indeed done with appliances which might to the ordinary spectator seem perplexingly crude, but when fully understood it was always clear that the conditions for the degree of accuracy aimed at were amply fulfilled. He had, in truth, no real liking for researches necessitating any extreme or meticulous refinement, he was too well aware what an expenditure of time these involve and was conscious of many more pressing things to be done. When a suggestion was made to amplify any line of his work with added detail, to carry the results to a higher degree of precision, he not infrequently advised "leaving it to the Germans." He was quite content with putting the problem upon the right lines and pursuing the conclusions up to the productive point; others could follow later into the arid regions of the third decimal place if they so desired and had no better use for their energies.

He was elected a Fellow in 1876, and served at different times in the offices of President of the Royal Astronomical, the Royal Photographic and the Physical Societies; also as Chairman of the Royal Society of Arts. He was President of Section A (Physics) at the British Asociation in 1889 where he gave an address summarising all that was known of the theory of photographic action. He was created K.C.B. in 1900. He married twice; first in 1864, Agnes Matilda, daughter of E. W. Smith, of Tickton Hall, Yorks, who died in 1888; secondly, in 1890, Mary Louisa, daughter of the Rev. G. N. Meade, of Scarborough on Hudson, U.S.A., who survives him.

He leaves one son and two daughters of the first marriage, one being the wife of Rear-Admiral Sir Reginald Hall; and one daughter of the second.

E. H. G-H.

ROBERT BELLAMY CLIFTON, 1836—1921.

Robert Bellamy Clifton was born at Gedney, in Lincolnshire, on March 13, 1836. He was the only son of Robert Clifton, a landowner in the district. At an early age he went to school at Peterborough, and afterwards at Brighton; then, after a period at University College, London, he entered at St. John's College, Cambridge, and took his degree as Sixth Wrangler in 1859. Canon J. M. Wilson, afterwards headmaster of Clifton, was Senior. Prof. Jack, who succeeded Clifton at Manchester, Prof. Adams, of King's College, London, and Dr. Stone, afterwards Radcliffe Observer, were in the same Tripos. Clifton was second Smith's Prizeman, Wilson being first, and it is clear from the testimonials he received when a candidate, in 1865, for the Professorship of Experimental Philosophy at Oxford, that his place in the Tripos was not thought by his contemporaries to represent his real merits.

He was elected a Fellow of his College shortly after taking his degree, and in 1860 became the first Professor of Natural Philosophy in Owen's College, Manchester. Roscoe was then Professor of Chemistry and in his 'Life and Experiences' writes:—"After a time R. B. Clifton, then a distinguished young Cambridge man, was appointed Professor of Physics, and this was the first step towards the expansion of the College in a scientific direction. Clifton soon became most popular; his lectures were admirable and enabled me to dispense with teaching any portion of his subject."

Clifton threw himself heartily into his work of teaching, giving annually two courses of experimental lectures, one course on elementary applied mathematics, with a short experimental course on the same subject, and occasional lectures involving higher mathematics. In his application for the Oxford chair, he states his conviction of the possibility of rendering intelligible to students possessing only the most elementary knowledge of mathematics, "by experimental lectures only, the results, both severally and in their mutual relations, of a thorough study of the science, provided the order in which the results are presented accords with that indicated by a strict mathematical investigation of the subject." This conviction he retained through life: it guided his work at Oxford, and the methods of instruction followed in the Clarendon Laboratory.

In 1865 he was appointed to the Professorship of Experimental Philosophy in the University of Oxford. The testimonials with which his application was supported form a striking collection. Stokes, Thomson, Adams, Joule, Roscoe, Bunsen, Kirchhoff, and Whewell write in the highest terms of his work and of the expectations formed of his ability and the hopes that in the freedom of Oxford it would find greater scope for research than had been possible at Manchester.

His first work in his new position was to design and build the Clarendon Laboratory, the first built in Europe for the special purpose of experimental instruction in Physics.

The money for its erection came from the proceeds of the sale of Clarendon's 'History of the Great Rebellion,' and in a letter Dr. Madan, formerly Bodley's Librarian, writes:

"Henry, Lord Cornbury, grandson of the historian, the first Earl of Clarendon, left many of the first earl's papers to trustees, with the direction that the money from the sale or publication of his papers should be the nucleus of a fund for an academy for riding or other exercises at Oxford. This was in 1751. But he died before his father. However, a sister carried out his intention, but the money was left to accumulate. In 1860 his trustees found they had £10,000, but as the University didn't need a riding school, but did badly want a laboratory for physical science, his trustees, by that wisdom which belongs only to lawyers and trustees, promptly erected the Clarendon Laboratory, which Prof. Clifton was the first to administer in 1872."

To design the Laboratory was his first task. The architect was responsible for the exterior; the fittings, down to minute details, were carried out from his own working drawings, and then, the building being complete, came its equipment with apparatus. Much of this was designed and redesigned by him until perfection, or something approaching it, was reached, and so much loving care had been spent on an instrument that it needed to be kept jealously under lock and key, taken out from time to time to be dusted and cleaned, possibly to be used in lecture, but entrusted never to the careless handling of a student of Physics.

At first he took some share in the practical instruction in the Laboratory, but for some time past this was left to his demonstrators. He was at the Laboratory every day, and generally went round and talked to the men, occasionally giving a demonstration on some special instrument, such as the Michelson interferometer.

Mechanics were outside his scope; until there was a special department for Electricity his lectures included a course on Electricity and Magnetism. After that they were chiefly concerned with Acoustics or Optics. His great idea was that his lectures should give his class instruction in giving experimental lectures; in the summer term he showed his men the solar spectrum in great detail with a grating spectroscope. He also lectured on the optical properties of crystals and the phenomena of polarisation, doing everything with minute care and great deliberation.

Of research work there was but little: the Laboratory was intended for teaching. This was the more unfortunate as the future careers of his pupils —among them may be mentioned Sir Arthur Rücker, Prof. Reinold, Sir Lazarus Fletcher, and others—have shown how competent they were to carry on research and advance knowledge, not merely by their teaching, but by original investigations.

Almost the only great piece of research appearing from the Clarendon Laboratory in his time was Boys' determination of the constant of gravitation, an investigation in which the author received the most cordial help and the most valuable support from the Professor.

To return to his own work. From 1879 to 1886 he was a member, along with Prof. Tyndall, Sir Frederick Abel, and others of the Royal Commission on Accidents in Mines, and assisted in carrying out a number of the experiments on which the conclusions of the Commission were based. He designed a safety lamp, which is described in Appendix XXIV of the Report and formed the subject of a number of the investigations. The general idea was to use the products of combustion so as to shield the flame from contact with an inflammable mixture.

He was elected a Fellow of the Royal Society in 1868, served three times on the Council, and was Vice-President in 1896–98. In 1869 he was elected a Fellow of Merton College, and at the time of his death he was an Honorary Fellow of Wadham. He published but little, partly because of his absorption in his teaching, but mainly, perhaps, as a consequence of his own high ideal of the level which published work should reach.

The 'Proceedings of the Literary and Philosophical Society of Manchester' for 1860–62 contain a paper, written jointly with Sir Henry Roscoe, on the "Effect of Increased Temperature upon the Nature of the Light emitted by the Vapours of certain Metals or Metallic Compounds," in which the suggestion, in the main correct, is made, that the flame spectra are due to the oxides of the metals, and not to the elements themselves. Another paper of interest, also published at Manchester, is one entitled "An Attempt to refer some Phenomena attending the Emission of Light to Mechanical Principles." It contains an early attempt to apply to the production of a spectrum some of the elementary facts of the kinetic theory of gases. His paper on "The Difference of Potential produced by the Contact of different Substances" ('Roy. Soc. Proc.,' vol. 26, 1877), was a useful contribution to the discussion then in progress as to the seat of the electromotive force in an electric circuit.

From early days he collected mathematical books of historical interest, building in later years a library, which he called the "Folly," on to his home at Oxford, to house these. He had a large collection of old Euclids, arithmetics, etc., and knew more about their contents than bibliophiles usually do. There is a story that, while still an undergraduate, Whewell sent for him and asked him his price for a copy of Calendri's 'De Arithmetica Opusculum' (Florence, 1491), which he had bought for 2*s.* 6*d.* Clifton declined to part with it, and Whewell, who wanted the book for a friend, said, "Well, I have done my duty by my friend, but now—you keep that book!"

Throughout his life he was a most kind friend to colleagues and pupils alike—a courteous gentleman, hospitable, and ready to help. A Manchester paper, referring to his successor in the Chair of Natural Philosophy, writes: "There is in him little of that easy grace which marked Mr. Clifton, making him most interesting and dignified when lecturing from a seat on the table, or when, having thrown away his gown, he worked with a vigour quite astounding to the audience at some laborious experiment"; the grace and dignity he retained to the end. According to "Who's Who," his recreation was "Work." The work included the supervision of a considerable property in Lincolnshire, which he had inherited from his father, and the welfare of his tenants, specially during a period of agricultural depression, which followed his father's death in 1873, was always uppermost in his thoughts. During term time, until he resigned his Professorship, he kept up a curious habit of dividing his day and night; he would rest in his chair after dinner for a time, and then when the household had gone to bed, he would start his work and go right on until 7.30 or 8 A.M. Then he went to bed for a couple of hours or so, and by 11 he was usually to be found at the Laboratory.

Prof. Clifton married, in 1862, Miss Butler, of Brighton, and leaves three sons and a daughter. Until 1893 their house at Oxford was always open to their friends, and they made a point of entertaining the men working at the Clarendon every term. In that year Mrs. Clifton had a serious illness, and for a time was cut off from all hospitality; before her death, in 1917, however, she recovered to a considerable extent, and the Sunday afternoon receptions again became quite an institution.

Oxford has changed; many of those who owed their training and their fortunes to Prof. Clifton have passed away, but the Clarendon Laboratory remains as a memorial to one who, though adopted, became her true and loyal son.

R. T. G.

SIR LAZARUS FLETCHER, 1854—1921.

SIR LAZARUS FLETCHER was born at Salford, March 3, 1854. He was the eldest of a family of six sons and two daughters, one of whom, the Rev. Mark Fletcher, F.G.S., is lecturer on Mineralogy at the Armstrong College, University of Durham, Newcastle-upon-Tyne.

He did not come from a scientific family, nor was it well-to-do; hence his attainments were not due to favourable circumstances in early life, but to his own love of work and learning, and the disadvantages which he had experienced made him very sympathetic with struggling students and ever ready to help them to the utmost of his power.

He was educated at the Manchester Grammar School under Dr. Marshall Watts, and for a time Mr. Francis Jones for Chemistry, Mr. Angell for Physics, and the Rev. J. Chambers for Mathematics, from whom he may have received his scientific bent as well as instruction. He changed over from the Science to the Mathematical VIth in 1871, and was equally good in both divisions, and while still at school he gained, in 1872, the Gold Medal for Mechanics and a Bronze Medal for Mathematics, and several first class certificates in other subjects at the National Examinations held by the Science and Art Department at South Kensington, also an open Science Scholarship (the Brackenbury) at Balliol, in 1872. He joined his College the same year and took a "highly distinguished" first class in Mathematical Mods. in 1874, and first classes in the Final Examinations in both the Natural Science and Mathematical Schools in 1876.

He was appointed Demonstrator in the Clarendon Physical Laboratory (1875–77) under Prof. R. B. Clifton, F.R.S., and was elected to the Millard Lectureship in Physics at Trinity College, Oxford, in 1877–78; also in 1877 to a Fellowship at University College, Oxford, which he had to resign in 1880 on account of his marriage, but was made an Hon. Fellow in 1910.

On account of his knowledge of Crystallography and Mineralogy he was appointed Assistant in the Mineralogy Department at the British Museum in 1878, under Prof. Nevil Story Maskelyne, F.R.S., in succession to Prof. W. G. Lewis, who had resigned, and he was promoted to the Keepership of the Department on Prof. Maskelyne's resignation in 1880, and held it till 1909.

Almost immediately after his appointment he had to carry out the removal of the great collections of Minerals and Meteorites from Bloomsbury and re-arrange them in the new buildings at South Kensington: probably the largest, finest and most valuable collection in the world. This was a most arduous and anxious task, not only on account of the vast number of specimens which had to be packed up for the journey and unpacked again and re-arranged, but also on account of the fragility of many, the very heavy weight of some, the great intrinsic value of the gold, diamonds and gems, and the priceless scientific value of the many rare and unique specimens, all of which, of course, had to be guarded against loss and theft, especially while out of their cases. He had to design new show-cases and alter others to adapt them to their new positions. Next he prepared Guides and Handbooks to these collections, which are models for simplicity and accuracy, and are helpful not only to the ordinary visitor but also to the student.

He was Examiner in Natural Science at the Oxford Public Examinations, 1880, and for the Cambridge Natural Science Tripos in 1882–3, 1889–91, 1896–7. He was elected a Fellow of the Royal Society in 1889, and served on the Council 1895, 1897, and 1910–12, and as a Vice-President in 1910–12, also a Vice-President of the Geological Society 1890–92; of the Physical Society 1895–7; President of the Mineralogical Society of Great Britain and Ireland 1885–8, and General Secretary 1888–1909; and President of the Geological Section, British Association, Oxford, 1894, and was awarded the Wollaston Medal of the Geological Society, 1912.

In presenting the Wollaston Medal of the Geological Society to Fletcher in 1912, the President, Prof. W. W. Watts, F.R.S., in the course of his speech, said "that the Council desired to place you in company with Bischof, Naumann, Dana, von Hauer, Descloizeaux, Story Maskelyne and von Groth.

. . . . I ask you to accept the Wollaston Medal, not for the sake of yourself and your work alone, but as a token of acknowledgement of the Science of Geology of part of her debt to the science which you so worthily represent in our country."

In the course of his acknowledgement, and in reference to the President's appreciative remarks upon his Optical Indicatrix paper, Fletcher said:—"After the work of Fresnel, Sir W. Hamilton, MacCullagh, Sylvester, and others, it was some time before he could convince himself that there was anything left for the gleaner, and that the relationship, if true, would have been discovered long ago." The relationship referred to is "that for every biaxial crystal to each point on the ellipsoid there corresponds a single ray of light, with three physical characters, viz., the direction and velocity of transmission and the plane of polarization—all definitely and simply related to the geometrical characters of the ellipsoid at that point."

He was appointed Director of the Natural History Branch of the British Museum in 1909, on the retirement of Sir Ray Lankester, K.C.B., was knighted in 1916, and retired in 1919.

It was during his directorship that the first guide-lecturer upon the Collections, Mr. T. H. Leonard, B.Sc., was appointed, and, in consequence, many visitors receive instruction which enables them to take an intelligent interest in what they see, instead of wandering aimlessly through the galleries and coming away with a museum headache. This innovation has been a most successful departure.

He was a Member of the Boards of Electors of Oxford and Cambridge to the Professorships of Mineralogy of both Universities. In addition to those already mentioned, he received numerous home and foreign honours, viz., he was made Corresponding Member of the Royal Society of Göttingen, the Academy of Sciences, Munich, and the New York Academy of Sciences. Hon. Member of the Soc. Científica, Antonio Alzate, Mexico; the Association Sci. et d'Enseignement Médical Complémentaire; the Selborne Society; the Hertfordshire Natural History Society and Field Club; the Ealing Science and Microscopical Society; and the Museums Association, and Hon. LL.D. of St. Andrews, and Hon. A.M. and Ph.D. of Berlin.

Over and above his administrative duties at the Natural History Museum, where he had larger collections to look after and a larger staff to supervise than when the Collections were at Bloomsbury, he found time for extra work in the way of research, and especially upon meteorites.

His services also to the Mineralogical Society were very great; he was President of it in 1885 to 1888, and its Hon. Secretary from 1888 to 1909; and, in recognition of the whole-hearted way he worked for it and ensured its success, the members and other friends subscribed for his portrait in oils and presented it to him in 1912.

The following are among other more important publications:—"An Introduction to the Study of Meteorites," 1881; "An Introduction to the Study of Minerals," 1884; "An Introduction to the Study of Rocks," 1895.

These were written as Guides to the Natural History Museum Collections, and ran through several editions. "On the Dilatation of Crystals," 'Phil. Mag.,' 1880; and that on "The Optical Indicatrix and Transmission of Light," 'Min. Mag.,' 1892, already referred to, in which he pointed out an error of Fresnel that had been copied by writers of text-books, etc., but is now corrected, thanks to Fletcher's investigations.

He delivered a notable Presidential Address to the Geological Section of the British Association—"On the Progress of Mineralogy and Crystallography," 1894, and also wrote the article on Meteorites in the 'Encyclopædia Britannica,' and prepared the 'Instructions for Collecting Rocks and Minerals in the Antarctic Manual,' 1901.

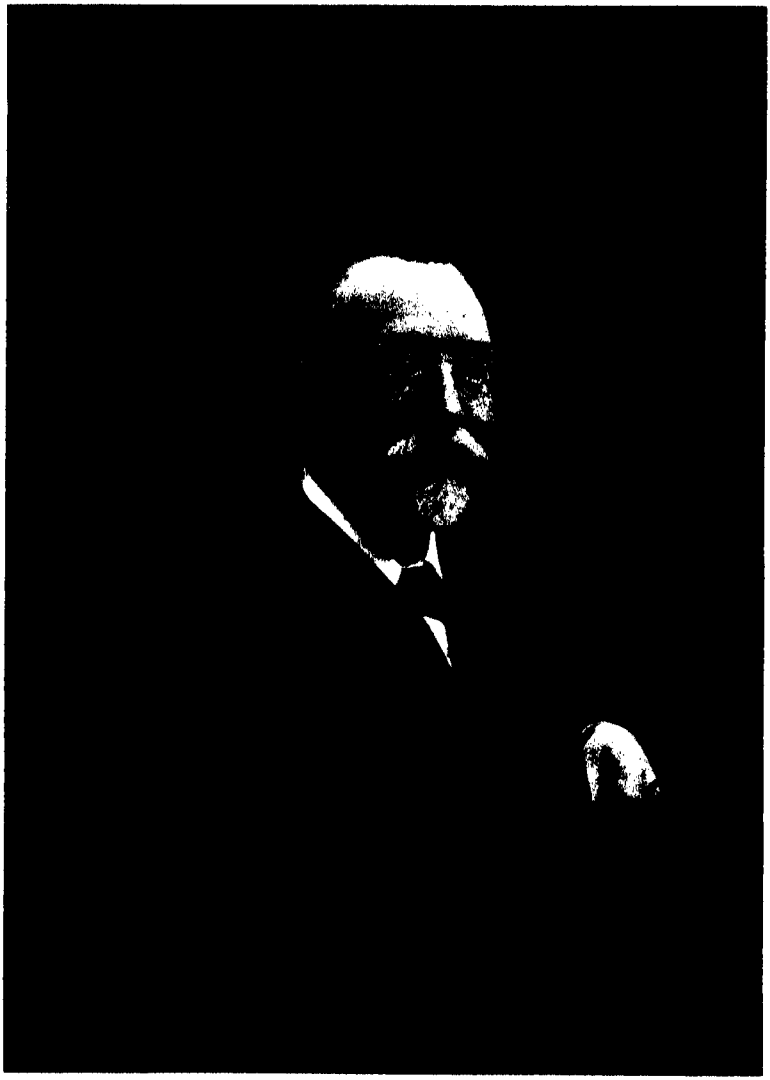
Most of his papers were read before the Mineralogical Society, and published in the 'Mineralogical Magazine,' and were upon Meteorites. He also gave a course of lectures on Meteorites in 1895, and again in 1904, at the Royal Institution.

Fletcher ever displayed great simplicity and charm of manner, and was of a most kindly and sympathetic disposition; the writer never heard him make a disparaging remark, or say an unkind word, about anyone; he always took a lenient view and gave credit for the best intentions; he was true to his friends, and would take any amount of trouble on their behalf; hence, for these and other reasons, he was held by them in affectionate regard.

He was gifted with humour, and would often sum up a matter by a whimsical humorous remark with a gentle smile and a merry twinkle, the piquancy of the sally often heightened by the assumption of a slight and pleasant Lancashire intonation. He was a model correspondent, even when far from well, and when unable to write himself he dictated, shortly before his death, quite a long letter to the present writer, with his good wishes for the New Year. He had a serious illness in 1906, and never really recovered his full strength and energy.

On his retirement from the Directorship he had to his own great regret and that of his friends, to resign his membership of the Athenæum and Savile Clubs, as his lessened income, the crushing income tax, and high cost of living in London, necessitated reduced expenditure; he therefore went to live at Ravenstonedale, a little village in Westmoreland. He was thus deprived of many social and other advantages, especially of the society of his friends, and of participation in the work of scientific societies and institutions.

He died at Grange-over-Sands on January 6 last, just before the date he had fixed for returning home, and was buried at Ravenstonedale on the 12th. On account of the distance but few of his many friends, to their regret, were able to attend his funeral. Lady Fletcher and his daughter survive him. A. L.



L. Fletcher

SRINIVASA RAMANUJAN, 1887—1920.

I.

SRINIVASA RAMANUJAN, who died at Kumbakonam on April 26, 1920, was elected a Fellow of the Society in 1918. He was not a man who talked much about himself, and until recently I knew very little of his early life. Two notices, by P. V. Seshu Aiyar and R. Ramachandra Rao, two of the most devoted of Ramanujan's Indian friends, have been published recently in the 'Journal of the Indian Mathematical Society'; and Sir Francis Spring has very kindly placed at my disposal an article which appeared in the 'Madras Times' of April 5, 1919. From these sources of information I can now supply a good many details with which I was previously unacquainted. Ramanujan (Srinivasa Iyengar Ramanuja Iyengar, to give him for once his proper name) was born on December 22, 1887, at Erode in southern India. His father was an accountant (*gumasta*) to a cloth merchant at Kumbakonam, while his maternal grandfather had served as *amin* in the Munsiff's (or local judge's) Court at Erode. He first went to school at five, and was transferred before he was seven to the Town High School at Kumbakonam, where he held a "free scholarship," and where his extraordinary powers appear to have been recognised immediately. "He used", so writes an old schoolfellow to Mr. Seshu Aiyar, "to borrow Carr's 'Synopsis of Pure Mathematics' from the College library, and delight in verifying some of the formulæ given there. . . . He used to entertain his friends with his theorems and formulæ, even in those early days. . . . He had an extraordinary memory, and could easily repeat the complete lists of Sanscrit roots (*atmanepada* and *parasmepada*); he could give the values of $\sqrt{2}$, π, e, ... to any number of decimal places. . . . In manners, he was simplicity itself. . . ."

He passed his matriculation examination to the Government College at Kumbakonam in 1904, and secured the "Junior Subraniam Scholarship". Owing to weakness in English, he failed in his next examination and lost his scholarship, and left Kumbakonam, first for Vizagapatam and then for Madras. Here he presented himself for the "First Examination in Arts" in December, 1906, but failed and never tried again. For the next few years he continued his independent work in mathematics, "jotting down his results in two good-sized note-books"; I have one of these note-books in my possession still. In 1909 he married, and it became necessary for him to find some permanent employment. I quote Mr. Seshu Aiyar:—

> To this end, he went to Tirukoilur, a small sub-division town in South Arcot District, to see Mr. V. Ramaswami Aiyar, the founder of the Indian Mathematical Society, but Mr. Aiyar, seeing his wonderful gifts, persuaded him to go to Madras. It was then after some four years' interval that Mr. Ramanujan met me at Madras, with his two

well-sized note-books referred to above. I sent Ramanujan with a note of recommendation to that true lover of mathematics, Dewan Bahadur R. Ramachandra Rao who was then District Collector at Nellore, a small town some eighty miles north of Madras. Mr. Rao sent him back to me saying it was cruel to make an intellectual giant like Ramanujan rot at a mofussil station like Nellore, and recommended his stay at Madras, generously undertaking to pay Mr. Ramanujan's expenses for a time. This was in December, 1910. After a while, other attempts to obtain for him a scholarship having failed, and Ramanujan himself being unwilling to be a burden on anybody for any length of time, he decided to take up a small appointment under the Madras Port Trust in 1911.

But he never slackened his work at mathematics. His earliest contribution to the 'Journal of the Indian Mathematical Society' was in the form of questions communicated by me in vol. 3 (1911). His first long article on "Some Properties of Bernoulli's Numbers" was published in the December number of the same volume. Mr. Ramanujan's methods were so terse and novel, and his presentation was so lacking in clearness and precision, that the ordinary reader, unaccustomed to such intellectual gymnastics, could hardly follow him. This particular article was returned more than once by the editor before it took a form suitable for publication. It was during this period that he came to me one day with some theorems on Prime Numbers, and when I referred him to Hardy's Tract on 'Orders of Infinity,' he observed that Hardy had said on p. 36 of his Tract "the exact order of $\rho(x)$ [defined by the equation

$$\rho(x) = \pi(x) - \int_2^x \frac{dt}{\log t},$$

where $\pi(x)$ denotes the number of primes less than x] has not yet been determined," and that he himself had discovered a result which gave the order of $\rho(x)$. On this I suggested that he might communicate his result to Mr. Hardy, together with some more of his results.

This passage brings me to the beginning of my own acquaintance with Ramanujan. But before I say anything about the letters which I received from him, and which resulted ultimately in his journey to England, I must add a little more about his Indian career. Dr. G. T. Walker, F.R.S., Head of the Meteorological Department, and formerly Fellow and Mathematical Lecturer of Trinity College, Cambridge, visited Madras for some official purpose some time in 1912, and Sir Francis Spring, K.C.I.E., the Chairman of the Madras Port Authority, called his attention to Ramanujan's work. Dr. Walker was far too good a mathematician not to recognise its quality, little as it had in common with his own. He brought Ramanujan's case to the notice of the Government and the University of Madras. A research studentship, "Rs. 75 *per mensem* for a period of two years", was awarded him, and he became, and remained for the rest of his life, a professional mathematician.

II.

Ramanujan wrote to me first on January 16, 1913, and at fairly regular intervals until he sailed for England in 1914. I do not believe that his letters were entirely his own. His knowledge of English, at that stage of his life, could scarcely have been sufficient, and there is an occasional phrase which is hardly characteristic. Indeed, I seem to remember his telling me that his friends had given him some assistance. However, it was the mathematics that mattered, and that was very emphatically his.

"Dear Sir, Madras, January 16, 1913.

"I beg to introduce myself to you as a clerk in the Accounts Department of the Port Trust Office at Madras on a salary of only £20 per annum. I am now about twenty-three years of age. I have have had no university education, but I have undergone the ordinary school course. After leaving school I have been employing the spare time at my disposal to work at mathematics. I have not trodden through the conventional regular course which is followed in a university course, but I am striking out a new path for myself. I have made a special investigation of divergent series in general, and the results I get are termed by the local mathematicians as 'startling'.

"Just as in elementary mathematics you give a meaning to a^n when n is negative and fractional to conform to the law which holds when n is a positive integer, similarly the whole of my investigations proceed on giving a meaning to Eulerian Second Integral for all values of n. My friends who have gone through the regular course of university education tell me that $\int_0^\infty x^{n-1}e^{-x}dx = \Gamma(n)$ is true only when n is positive. They say that this integral relation is not true when n is negative. Supposing this is true only for positive values of n, and also supposing the definition $n\Gamma(n) = \Gamma(n+1)$ to be universally true, I have given meanings to these integrals, and under the conditions I state the integral is true for all values of n negative and fractional. My whole investigations are based upon this, and I have been developing this to a remarkable extent, so much so that the local mathematicians are not able to understand me in my higher flights.

"Very recently I came across a tract published by you styled 'Orders of Infinity', in p. 36 of which I find a statement that no definite expression has been as yet found for the no of prime nos less than any given number. I have found an expression which very nearly approximates to the real result, the error being negligible. I would request you to go through the enclosed papers. Being poor, if you are convinced that there is anything of value, I would like to have my theorems published. I have not given the actual investigations nor the expressions that I get, but I have indicated to the lines on which I proceed. Being inexperienced, I would very highly value any advice you give me. Requesting to be excused for the trouble I give you.

"I remain, Dear sir, Yours truly,
"S. RAMANUJAN.

"P.S.—My address is S. Ramanujan, Clerk Accounts Department, Port Trust, Madras, India."

I quote now from the "papers enclosed," and from later letters.

"The following are a few examples from my theorems :—

(1) The nos of the form $2^p 3^q$ less than $n = \frac{1}{2} \frac{\log (2n) \log (3n)}{\log 2 \log 3}$ where p and q may have any positive integral value including 0.

(2) Let us take all nos containing an odd no of dissimilar prime divisors, viz. :—

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 30, 31, 37, 41, 42, 43, 47, etc.

(*a*) The no of such nos less than $n = \frac{3n}{\pi^2}$.

(*b*) $\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{5^2}+\frac{1}{7^2}+\ldots+\frac{1}{30^2}+\frac{1}{31^2}+\ldots = \frac{9}{2\pi^2}$.

(*c*) $\frac{1}{2^4}+\frac{1}{3^4}+\frac{1}{5^4}+\frac{1}{7^4}+$ etc. $= \frac{15}{2\pi^4}$.

(3) Let us take the no of divisors of natural nos, viz. :—

1, 2, 2, 3, 2, 4, 2, 4, 3, 4, 2, etc. (1 having 1 divisor, 2 having 2, 3 having 2, 4 having 3, 5 having 2, etc.).

The sum of such nos to n terms

$$= n(2\gamma - 1 + \log n) + \tfrac{1}{2} \text{ of the no of divisors of } n,$$

where $\gamma = 0{\cdot}5772156649 \ldots$, the Eulerian Constant.

(4) 1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, etc., are nos which are either themselves sqq. or which can be expressed as the sum of two sqq.

"The no of such nos greater than A and less than B

$$= K \int_A^B \frac{dx}{\sqrt{\log x}} + \theta(B) \quad \text{where} \quad K = 0{\cdot}764 \ldots$$

and $\theta(x)$ is very small when compared with the previous integral. K and $\theta(x)$ have been exactly found though complicated. . . ."

It may be well that I should interpolate here a few remarks concerning Ramanujan's researches in this particular field.

Ramanujan's theory of primes was vitiated by his ignorance of the theory of functions of a complex variable. It was (so to say) what the theory might be if the Zeta-function had no complex zeros. His methods of proof depended upon a wholesale use of divergent series. He disregarded entirely all the difficulties which are involved in the interchange of double limit operations; he did not distinguish, for example, between the sum of a series Σa_n and the value of the Abelian limit

$$\lim_{x \to 1} \Sigma a_n x^n,$$

or that of any other limit which might be used for similar purposes by a modern analyst. There are regions of mathematics in which the precepts of modern rigour may be disregarded with comparative safety, but the Analytic Theory of Numbers is not one of them, and Ramanujan's Indian work on primes, and on all the allied problems of the theory, was definitely wrong. That his proofs should have been invalid was only to be expected. But the mistakes went deeper than that, and many of the actual results were false. He had obtained the dominant terms of the classical formulæ, although by invalid methods; but none of them are such close approximations as he supposed.

This may be said to have been Ramanujan's one great failure. And yet I am not sure that, in some ways, his failure was not more wonderful than any of his triumphs. Consider, for example, problem (4). The dominant term is

$$\frac{KB}{\sqrt{\log B}} \tag{α}$$

(I adhere to Ramanujan's notation): this result was first obtained by Landau in 1908. The error is of order $B(\log B)^{-\frac{3}{2}}$; and this is so far in agreement with Ramanujan's assertion. Ramanujan, however, implies much more,* and what he implies is definitely false: his integral does not represent the number of numbers in question more accurately than Landau's formula (α). However, Ramanujan had none of Landau's weapons at his command; he had never seen a French or German book; his knowledge even of English was insufficient to enable him to qualify for a degree. It is sufficiently marvellous that he should have even dreamt of problems such as these, problems which it has taken the finest mathematicians in Europe a hundred years to solve, and of which the solution is incomplete to the present day.

"... IV. Theorems on integrals. The following are a few examples :—

(1) $$\int_0^\infty \frac{1+\left(\dfrac{x}{b+1}\right)^2}{1+\left(\dfrac{x}{a}\right)^2} \cdot \frac{1+\left(\dfrac{x}{b+2}\right)^2}{1+\left(\dfrac{x}{a+1}\right)^2} \ldots \text{etc. } dx$$

$$= \frac{\sqrt{\pi}}{2} \cdot \frac{\Gamma(a+\frac{1}{2})}{\Gamma(a)} \cdot \frac{\Gamma(b+1)}{\Gamma(b+\frac{1}{2})} \cdot \frac{\Gamma(b-a+\frac{1}{2})}{\Gamma(b-a+1)}.$$

......

(3) If $$\int_0^\infty \frac{\cos nx}{e^{2\pi\sqrt{x}}-1}\,dx = \phi(n),$$

then $$\int_0^\infty \frac{\sin nx}{e^{2\pi\sqrt{x}}-1}\,dx = \phi(n) - \frac{1}{2n} + \phi\left(\frac{\pi^2}{n}\right)\sqrt{\frac{2\pi^3}{n^3}}.$$

---

* See his statement concerning the order of $\theta(x)$, quoted below from his letter of February 27, 1913.

$\phi(n)$ is a complicated function. The following are certain special values:

$$\phi(0)=\frac{1}{12};\ \phi\left(\frac{\pi}{2}\right)=\frac{1}{4\pi};\ \ \phi(\pi)=\frac{2-\sqrt{2}}{8};\ \phi(2\pi)=\frac{1}{16};$$

$$\phi\left(\frac{2\pi}{5}\right)=\frac{8-8\sqrt{5}}{16};\ \phi\left(\frac{\pi}{5}\right)=\frac{6+\sqrt{5}}{4}-\frac{5\sqrt{10}}{8};\ \phi(\infty)=0;$$

$$\phi\left(\frac{2\pi}{3}\right)=\frac{1}{3}-\sqrt{3}\left(\frac{3}{16}-\frac{1}{8\pi}\right).$$

(4) $$\int_0^\infty \frac{dx}{(1+x^2)(1+r^2x^2)(1+r^4x^2)\ldots\text{etc.}}=\frac{\pi}{2(1+r+r^3+r^6+r^{10}+\text{etc.})},$$

where 1, 3, 6, 10, etc., are sums of natural nos.

……

(6) $\displaystyle\int_0^\infty \frac{\tan^{-1}\dfrac{2nz}{n^2+x^2-z^2}}{e^{2\pi z}-1}\,dz$ can be exactly found if $2n$ is any integer and x any quantity."

……

"V. Theorems on summation of series;[*] *e.g.*,

(1) $$\frac{1}{1^3}\cdot\frac{1}{2}+\frac{1}{2^3}\cdot\frac{1}{2^2}+\frac{1}{3^3}\cdot\frac{1}{2^3}+\frac{1}{4^3}\cdot\frac{1}{2^4}+\text{etc.}$$

$$=\tfrac{1}{6}(\log 2)^3-\frac{\pi^2}{12}\log 2+\left(\frac{1}{1^3}+\frac{1}{3^3}+\frac{1}{5^3}+\text{etc.}\right).$$

(2) $$1+9\,.\,(\tfrac{1}{4})^4+17\,.\left(\frac{1\,.\,5}{4\,.\,8}\right)^4+25\,.\left(\frac{1\,.\,5\,.\,9}{4\,.\,8\,.\,12}\right)^4+\text{etc.}=\frac{2\sqrt{2}}{\sqrt{\pi}\,.\,\{\Gamma(\frac{3}{4})\}^2}.$$

(3) $$1-5\,.\,(\tfrac{1}{2})^3+9\,.\left(\frac{1\,.\,3}{2\,.\,4}\right)^3-\text{etc.}=\frac{2}{\pi}.$$

(4) $$\frac{1^{13}}{e^{2\pi}-1}+\frac{2^{13}}{e^{4\pi}-1}+\frac{3^{13}}{e^{6\pi}-1}+\text{etc.}=\frac{1}{24}.$$

(5) $$\frac{\coth\pi}{1^7}+\frac{\coth 2\pi}{2^7}+\frac{\coth 3\pi}{3^7}+\text{etc.}=\frac{19\pi^7}{56700}.$$

……

(11) $$\frac{2}{3}\int_0^1\frac{\tan^{-1}x}{x}dx-\int_0^{2-\sqrt{3}}\frac{\tan^{-1}x}{x}dx=\frac{\pi}{12}\log(2+\sqrt{3}).$$

……

[*] There is always more in one of Ramanujan's formulæ than meets the eye, as anyone who sets to work to verify those which look the easiest will soon discover. In some the interest lies very deep, in others comparatively near the surface; but there is not one which is not curious and entertaining.

" VI. Theorems on transformation of series and integrals, *e.g.*,

(1) $$\pi\left(\frac{1}{2}-\frac{1}{\sqrt{1}+\sqrt{3}}+\frac{1}{\sqrt{3}+\sqrt{5}}-\frac{1}{\sqrt{5}+\sqrt{7}}+\text{ etc.}\right)=\frac{1}{1\sqrt{1}}-\frac{1}{3\sqrt{3}}+\frac{1}{5\sqrt{5}}-\text{ etc.}$$

......

(4) If $\int_0^a \phi(p, x)\cos nx\,dx = \psi(p, n)$, then

$$\frac{\pi}{2}\int_0^\infty \phi(p, x)\,\phi(q, nx)\,dx = \int_0^\infty \phi(q, x)\,\phi(p, nx)\,dx.$$

(5) If $\alpha\beta = \pi$, then $$\sqrt{\alpha}\int_0^\infty \frac{e^{-x^2}}{\cosh \alpha x}\,dx = \sqrt{\beta}\int_0^\infty \frac{e^{-x^2}}{\cosh \beta x}\,dx.$$

......

" VII. Theorems on approximate integration and summation of series.

(1) $1^2\log 1+2^2\log 2+3^2\log 3+\ldots+x^2\log x$

$$=\frac{x(x+1)(2x+1)}{6}\log x-\frac{x^3}{9}+\frac{1}{4\pi^2}\left(\frac{1}{1^3}+\frac{1}{3^3}+\ldots\right)+\frac{x}{12}-\frac{1}{360\,x}+\ldots\ldots$$

(2) $$1+\frac{x}{\underline{|1}}+\frac{x^2}{\underline{|2}}+\frac{x^3}{\underline{|3}}+\ldots+\frac{x^x}{\underline{|x}}\,\theta=\frac{e^x}{2}$$

where $\theta=\frac{1}{3}+\frac{4}{135\,(x+k)}$ where k lies between $\frac{8}{45}$ and $\frac{2}{21}$.

(3) $$1+\left(\frac{x}{\underline{|1}}\right)^5+\left(\frac{x^2}{\underline{|2}}\right)^5+\left(\frac{x^3}{\underline{|3}}\right)^5+\text{ etc.}=\frac{\sqrt{5}}{4\pi^2}\cdot\frac{e^{5x}}{5x^2-x+\theta},$$

where θ vanishes when $x=\infty$.

......

(5) $$\frac{1}{1001}+\frac{1}{1002^2}+\frac{3}{1003^3}+\frac{4^2}{1004^4}+\frac{5^3}{1005^5}+\text{ etc.}=\frac{1}{1000}-10^{-440}\times 1{\cdot}0125\text{ nearly.}$$

(6) $$\int_0^a e^{-x^2}\,dx=\frac{\sqrt{\pi}}{2}-\frac{e^{-a^2}}{2a}\,{}_{+}\,\frac{1}{a}{}_{+}\frac{2}{2a}{}_{+}\frac{3}{a}{}_{+}\frac{4}{2a}{}_{+}\text{etc.}$$

(7) The coefficient of x^n in $\frac{1}{1-2x+2x^4-2x^9+2x^{16}-\text{etc.}}$

$$=\text{the nearest integer to }\frac{1}{4n}\left\{\cosh(\pi\sqrt{n})-\frac{\sinh(\pi\sqrt{n})}{\pi\sqrt{n}}\right\}.^*$$

---

* **This is quite untrue. But the formula is extremely interesting for a variety of reasons.**

"IX. Theorems on continued fractions, a few examples are:—

(1) $$\frac{4}{x+}\,\frac{1^2}{2x+}\,\frac{3^2}{2x+}\,\frac{5^2}{2x+}\,\frac{7^2}{2x+}\text{ etc.}=\left\{\frac{\Gamma\left(\frac{x+1}{4}\right)}{\Gamma\left(\frac{x+3}{4}\right)}\right\}^2.$$

......

(4) If $$u=\frac{x}{1+}\,\frac{x^5}{1+}\,\frac{x^{10}}{1+}\,\frac{x^{15}}{1+}\,\frac{x^{20}}{1+}\text{ etc.}$$

and $$v=\frac{\sqrt[5]{x}}{1}\,\frac{x}{+1+}\,\frac{x^2}{1+}\,\frac{x^3}{1+}\text{ etc.}$$

then $$v^5=u\,.\,\frac{1-2u+4u^2-3u^3+u^4}{1+3u+4u^2+2u^3+u^4}.$$

(5) $$\frac{1}{1+}\,\frac{e^{-2\pi}}{1+}\,\frac{e^{-4\pi}}{1+}\,\frac{e^{-6\pi}}{1+}\text{ etc.}=\left(\sqrt{\frac{5+\sqrt{5}}{2}}-\frac{\sqrt{5}+1}{2}\right)\sqrt[5]{e^{2\pi}}.$$

(6) $$\frac{1}{1-}\,\frac{e^{-\pi}}{1+}\,\frac{e^{-2\pi}}{1-}\,\frac{e^{-3\pi}}{1+}\text{ etc.}=\left(\sqrt{\frac{5-\sqrt{5}}{2}}-\frac{\sqrt{5}-1}{2}\right)\sqrt[5]{e^{\pi}}.$$

(7) $\frac{1}{1+}\,\frac{e^{-\pi\sqrt{n}}}{1+}\,\frac{e^{-2\pi\sqrt{n}}}{1+}\,\frac{e^{-3\pi\sqrt{n}}}{1+}\text{ etc.}$ can be exactly found if n be any positive rational quantity. . . ."

February 27, 1913.

". . . I have found a friend in you who views my labours sympathetically. This is already some encouragement to me to proceed. . . . I find in many a place in your letter rigorous proofs are required, and you ask me to communicate the methods of proof. . . . I told him* that the sum of an infinite no of terms of the series $1+2+3+4+\ldots=-\frac{1}{12}$ under my theory. If I tell you this you will at once point out to me the lunatic asylum as my goal. . . What I tell you is this. Verify the results I give, and, if they agree with your results . . . you should at least grant that there may be some truths in my fundamental basis. . . .

"To preserve my brains I want food, and this is now my first consideration. Any sympathetic letter from you will be helpful to me here to get a scholarship, either from the University or from Government. . . .

1. The no of prime nos less than $e^a=\int_0^\infty\frac{a^x dx}{xS_{x+1}\Gamma(x+1)}$

where $$S_{x+1}=\frac{1}{1^{x+1}}+\frac{1}{2^{x+1}}+\ldots$$

---

* Referring to a previous correspondence.

2. The no of prime nos less than n =

$$\frac{2}{\pi}\left\{\frac{2}{B_2}\left(\frac{\log n}{2\pi}\right)+\frac{4}{3B_4}\left(\frac{\log n}{2\pi}\right)^3+\frac{6}{5B_6}\left(\frac{\log n}{2\pi}\right)^5+\text{ etc.}\right.$$

where $B_2 = \frac{1}{6}$; $B_4 = \frac{1}{30}$, etc., the Bernoullian nos. . . .

……

The order of $\theta(x)$ which you asked in your letter is $\sqrt{\left(\dfrac{x}{\log x}\right)}$.

……

(1) If $F(x) = \dfrac{1}{1+}\dfrac{x}{1+}\dfrac{x^2}{1+}\dfrac{x^3}{1+}\dfrac{x^4}{1+}\dfrac{x^5}{1+}$ etc.

then $$\left\{\frac{\sqrt{5}+1}{2}+e^{-2\alpha/5}F(e^{-2\alpha})\right\}\left\{\frac{\sqrt{5}+1}{2}+e^{-2\beta/5}F(e^{-2\beta})\right\}=\frac{5+\sqrt{5}}{2},$$

with the condition $\alpha\beta = \pi^2$. . . .

e.g. $$\frac{1}{1+}\frac{e^{-2\pi\sqrt{5}}}{1+}\frac{e^{-4\pi\sqrt{5}}}{1}+\text{ etc. }\ldots=e^{2\pi/\sqrt{5}}\left\{\frac{\sqrt{5}}{1+\sqrt[5]{5^{\frac{3}{4}}\left(\frac{\sqrt{5}-1}{2}\right)^{\frac{5}{2}}-1}}-\frac{\sqrt{5}+1}{2}\right\}$$

The above theorem is a particular case of a theorem on the c.f.

$$\frac{1}{1+}\frac{ax}{1+}\frac{ax^2}{1+}\frac{ax^3}{1+}\frac{ax^4}{1+}\frac{ax^5}{1+}\text{ etc.}$$

which is a particular case of the c.f.

$$\frac{1}{1+}\frac{ax}{1+bx+}\frac{ax^2}{1+bx^2+}\frac{ax^3}{1+bx^3+}\text{ etc.}$$

which is a particular case of a general theorem on c.f.

(2) i. $$4\int_0^\infty\frac{xe^{-x\sqrt{5}}}{\cosh x}dx=\frac{1}{1+}\frac{1^2}{1+}\frac{1^2}{1+}\frac{2^2}{1+}\frac{2^2}{1+}\frac{3^2}{1+}\frac{3^2}{1+}\text{ etc.}$$

ii. $$4\int_0^\infty\frac{x^2e^{-x\sqrt{3}}}{\cosh x}dx=\frac{1}{1+}\frac{1^3}{1+}\frac{1^3}{3+}\frac{2^3}{1+}\frac{2^3}{5+}\frac{3^3}{1+}\frac{3^3}{7+}\text{ etc.}$$

……

(5) $$\frac{1^5}{e^{2\pi}-1}\frac{1}{2500+1^4}+\frac{2^5}{e^{4\pi}-1}\frac{1}{2500+2^4}+\ldots=\frac{123826979}{6306456}-\frac{25\pi}{4}\coth^2 5\pi.$$

(6) If $v=\dfrac{x}{1+}\dfrac{x^3+x^6}{1+}\dfrac{x^6+x^{12}}{1+}\dfrac{x^9+x^{18}}{1+}$ etc.

then i. $$x\left(1+\frac{1}{v}\right)=\frac{1+x+x^3+x^6+x^{10}+\text{ etc.}}{1+x^9+x^{27}+x^{54}+x^{90}+\text{ etc.}}$$

ii. $$x^3\left(1+\frac{1}{v^3}\right)=\left(\frac{1+x+x^3+x^6+{}^{10}+\text{ etc.}}{1+x^3+x^9+x^{18}+x^{30}+\text{ etc.}}\right)^4.$$

……

(8) $$\int_0^\infty \frac{1+ab^2x^2}{1+x^2}\,\frac{1+ab^4x^2}{1+b^2x^2}\dots\,dx = \frac{\pi}{2(1+b+b^3+b^6+\dots)}\,\frac{1-ab^2}{1-ab}\,\frac{1-ab^4}{1-ab^3}\dots$$

……

(12) If $\frac{1}{2}\pi\alpha = \log\tan(\frac{1}{4}\pi+\frac{1}{4}\pi\beta)$, then

$$\frac{1^2+\alpha^2}{1^2-\beta^2}\left(\frac{3^2-\beta^2}{3^2+\alpha^2}\right)^3\left(\frac{5^2+\alpha^2}{5^2-\beta^2}\right)^5\left(\frac{7^2-\beta^2}{7^2+\alpha^2}\right)^7\dots = e^{\frac{1}{2}\pi\alpha\beta}.$$

……

(16) If $F(\alpha, \beta) = \tan^{-1}\left(\frac{\alpha}{x+}\,\frac{\beta^2+k^2}{x+}\,\frac{\alpha^2+(2k)^2}{x+}\,\frac{\beta^2+(3k)^2}{x}\dots\right)$, then

$$F(\alpha, \beta)+F(\beta, \alpha) = 2F\left(\frac{\alpha+\beta}{2}, \frac{\alpha+\beta}{2}\right).$$

(17) If $F(k) = 1+\left(\frac{1}{2}\right)^2 k+\left(\frac{1\,.\,3}{2\,.\,4}\right)^2 k^2+\dots$ and $F(1-k) = \sqrt{(210)}\,F(k)$,

then $k = (\sqrt{2}-1)^4(2-\sqrt{3})^2(\sqrt{7}-\sqrt{6})^4(8-3\sqrt{7})^2(\sqrt{10}-3)^4(4-\sqrt{15})^4$
$\times(\sqrt{15}-\sqrt{14})^2(6-\sqrt{35})^2$.

(18) If $$F(\alpha) = 1+\frac{1\,.\,2}{3^2}\alpha+\frac{1\,.\,2\,.\,4\,.\,5}{3^2\,.\,6^2}\alpha^2+\dots$$

and $$\frac{F(1-\alpha)}{F(\alpha)} = 5\frac{F(1-\beta)}{F(\beta)},$$

then $$\sqrt[3]{(\alpha\beta)}+\sqrt[3]{\{(1-\alpha)(1-\beta)\}}+3\sqrt[6]{\{\alpha\beta(1-\alpha)(1-\beta)\}} = 1.$$

……

April 17, 1913.

". . . I am a little pained to see what you have written. . . .* I am not in the least apprehensive of my method being utilised by others. On the contrary, my method has been in my possession for the last eight years, and I have not found anyone to appreciate the method. As I wrote in my last letter, I have found a sympathetic friend in you, and I am willing to place unreservedly in your hands what little I have. It was on account of the novelty of the method I have used that I am a little diffident even now to communicate my own way of arriving at the expressions I have already given. . . .

". . . I am glad to inform you that the local University has been pleased to grant me a scholarship of £60 per annum for two years, and this was at the instance of Dr. Walker, F.R.S., Head of the Meteorological Department in India, to whom my thanks are due. . . . I request you to convey my thanks also to Mr. Littlewood, Dr. Barnes, Mr. Berry, and others who take an interest in me. . . ."

---

* Ramanujan might very reasonably have been reluctant to give away his secrets to an English mathematician, and I had tried to reassure him on this point as well as I could.

III.

It is unnecessary to repeat the story of how Ramanujan was brought to England. There were serious difficulties; and the credit for overcoming them is due primarily to Prof. E. H. Neville, in whose company Ramanujan arrived in April, 1914. He had a scholarship from Madras of £250, of which £50 was allotted to the support of his family in India, and an exhibition of £60 from Trinity. For a man of his almost ludicrously simple tastes, this was an ample income; and he was able to save a good deal of money, which was badly wanted later. He had no duties, and could do as he pleased; he wished, indeed, to qualify for a Cambridge degree as a research student, but this was a formality. He was now, for the first time in his life, in a really comfortable position, and could devote himself to his researches without anxiety.

There was one great puzzle. What was to be done in the way of teaching him modern mathematics? The limitations of his knowledge were as startling as its profundity. Here was a man who could work out modular equations, and theorems of complex multiplication, to orders unheard of, whose mastery of continued fractions was, on the formal side at any rate, beyond that of any mathematician in the world, who had found for himself the functional equation of the Zeta-function, and the dominant terms of many of the most famous problems in the analytic theory of numbers; and he had never heard of a doubly periodic function or of Cauchy's theorem, and had, indeed, but the vaguest idea of what a function of a complex variable was. His ideas as to what constituted a mathematical proof were of the most shadowy description. All his results, new or old, right or wrong, had been arrived at by a process of mingled argument, intuition, and induction, of which he was entirely unable to give any coherent account.

It was impossible to ask such a man to submit to systematic instruction, to try to learn mathematics from the beginning once more. I was afraid, too, that, if I insisted unduly on matters which Ramanujan found irksome, I might destroy his confidence or break the spell of his inspiration. On the other hand, there were things of which it was impossible that he should remain in ignorance. Some of his results were wrong, and in particular those which concerned the distribution of primes, to which he attached the greatest importance. It was impossible to allow him to go through life supposing that all the zeros of the Zeta-function were real. So I had to try to teach him, and in a measure I succeeded, though obviously I learnt from him much more than he learnt from me. In a few years' time he had a very tolerable knowledge of the theory of functions and the analytic theory of numbers. He was never a mathematician of the modern school, and it was hardly desirable that he should become one; but he knew when he had proved a theorem and when he had not. And his flow of original ideas showed no symptom of abatement.

I should add here a word about Ramanujan's interests outside mathematics. Like his mathematics, they showed the strangest contrasts. He had very little interest, I should say, in literature as such, or in art, though he could tell good literature from bad. On the other hand, he was a keen philosopher, of what appeared, to followers of the modern Cambridge school, a rather nebulous kind, and an ardent politician, of a pacifist and ultra-radical type. He adhered, with a severity most unusual in Indians resident in England, to the religious observances of his caste; but his religion was a matter of observance and not of intellectual conviction, and I remember well his telling me (much to my surprise) that all religions seemed to him more or less equally true. Alike in literature, philosophy, and mathematics, he had a passion for what was unexpected, strange, and odd; he had quite a small library of books by circle-squarers and other cranks.

It was in the spring of 1917 that Ramanujan first appeared to be unwell. He went into the Nursing Home at Cambridge in the early summer, and was never out of bed for any length of time again. He was in sanatoria at Wells, at Matlock, and in London, and it was not until the autumn of 1918 that he showed any decided symptom of improvement. He had then resumed active work, stimulated perhaps by his election to the Royal Society, and some of his most beautiful theorems were discovered about this time. His election to a Trinity Fellowship was a further encouragement. Early in 1919 he had recovered, it seemed, sufficiently for the voyage home to India, and the best medical opinion held out hopes of a permanent restoration. I was rather alarmed by not hearing from him for a considerable time; but a letter reached me in February, 1920, from which it appeared that he was still active in research.

University of Madras,
January 12, 1920.

"I am extremely sorry for not writing you a single letter up to now. . . . I discovered very interesting functions recently which I call 'Mock' ϑ-functions. Unlike the 'False' ϑ-functions (studied partially by Prof. Rogers in his interesting paper), they enter into mathematics as beautifully as the ordinary ϑ-functions. I am sending you with this letter some examples. . . .

Mock ϑ-functions

$$\phi(q) = 1 + \frac{q}{1+q^2} + \frac{q^4}{(1+q^2)(1+q^4)} + \dots$$

$$\psi(q) = \frac{q}{1-q} + \frac{q^4}{(1-q)(1-q^3)} + \frac{q^9}{(1-q)(1-q^3)(1-q^5)} + \dots$$

Mock ϑ-functions (of 5th order)

$$f(q) = 1 + \frac{q}{1+q} + \frac{q^4}{(1+q)(1+q^2)} + \frac{q^9}{(1+q)(1+q)(1+q^3)} + \dots$$

"*Mock ϑ-functions (of 7th order)*

$$\text{(i)}\quad 1+\frac{q}{1-q^2}+\frac{q^4}{(1-q^3)(1-q^4)}+\frac{q^9}{(1-q^4)(1-q^5)(1-q^6)}+\dots$$

……"

He said little about his health, and what he said was not particularly discouraging; and I was quite unprepared for the news of his death.

IV.

Ramanujan published the following papers in Europe:—

(1) "Some Definite Integrals," 'Messenger of Mathematics,' vol. 44, pp. 10–18 (1914).
(2) "Some Definite Integrals connected with Gauss's sums," *ibid.*, pp. 75–85.
(3) "Modular Equations and Approximations to π," 'Quarterly Journal of Mathematics,' vol. 45, pp. 350–372 (1914).
(4) "New Expressions for Riemann's Functions $\zeta(s)$ and $\Xi(t)$," *ibid.*, vol. 46, pp. 253–261 (1915).
(5) "On Certain Infinite Series," 'Messenger of Mathematics,' vol. 45, pp. 11–15 (1915).
(6) "Summation of a Certain Series," *ibid.*, pp. 157–160.
(7) "Highly Composite Numbers," 'Proc. London Math. Soc.,' ser. 2, vol. 14, pp. 347–409 (1915).
(8) "Some Formulæ in the Analytic Theory of Numbers," 'Messenger of Mathematics,' vol. 45, pp. 81–84 (1916).
(9) "On Certain Arithmetical Functions," 'Trans. Cambridge Phil. Soc.,' vol. 22, No. 9, pp. 159–184 (1916).
(10) "Some Series for Euler's Constant," 'Messenger of Mathematics,' vol. 46, pp. 73–80 (1916).
(11) "On the Expression of Numbers in the Form $ax^2+by^2+cz^2+dt^2$," 'Proc. Cambridge Phil. Soc.,' vol. 19, pp. 11–21 (1917).
*(12) "Une formule asymptotique pour le nombre des partitions de n," 'Comptes Rendus,' January 2, 1917.
*(13) "Asymptotic Formulæ concerning the Distribution of Integers of various Types," 'Proc. London Math. Soc.,' ser. 2, vol. 16, pp. 112–132 (1917).
*(14) "The Normal Number of Prime Factors of a Number n," 'Quarterly Journal of Mathematics,' vol. 48, pp. 76–92 (1917).
*(15) "Asymptotic Formulæ in Combinatory Analysis," 'Proc. London Math. Soc.,' ser. 2, vol. 17, pp. 75–115 (1918).
*(16) "On the Coefficients in the Expansions of Certain Modular Functions," 'Roy. Soc. Proc.,' A, vol. 95, pp. 144–155 (1918).
(17) "On Certain Trigonometrical Sums and their Applications in the Theory of Numbers," 'Trans. Cambridge Phil. Soc.,' vol. 22, pp. 259–276 (1918).
(18) "Some Properties of $p(n)$, the Number of Partitions of n," 'Proc. Cambridge Phil. Soc.,' vol. 19, pp. 207–210 (1919).
(19) "Proof of Certain Identities in Combinatory Analysis," *ibid.*, pp. 214–216.
(20) "A Class of Definite Integrals," 'Quarterly Journal of Mathematics,' vol. 48, pp. 294–309 (1920).
(21) "Congruence Properties of Partitions," 'Math. Zeitschrift,' vol. 9, pp. 147–153 (1921).

Of these, those marked with an asterisk were written in collaboration with me, and (21) is a posthumous extract from a much larger unpublished manuscript in my possession.† He also published a number of short notes in the 'Records of Proceedings at Meetings' of the London Mathematical Society, and in the 'Journal of the Indian Mathematical Society.' The complete list of these is as follows :—

Records of Proceedings at Meetings.

*(22) "Proof that almost all Numbers n are Composed of about log log n prime factors," December 14, 1916.
*(23) "Asymptotic Formulæ in Combinatory Analysis," March 1, 1917.
(24) "Some Definite Integrals," Jan. 17, 1918.
(25) "Congruence Properties of Partitions," March 13, 1919.
(26) "Algebraic Relations between certain Infinite Products," March 13, 1919.

Journal of the Indian Mathematical Society.

(A) Articles and Notes.

(27) "Some Properties of Bernoulli's Numbers," vol. 3, pp. 219–235 (1911).
(28) "On Q. 330 of Prof. Sanjaná," vol. 4, pp. 59–61 (1912).
(29) "A Set of Equations," vol. 4, pp. 94–96 (1912).
(30) "Irregular Numbers," vol. 5, pp. 105–107 (1913).
(31) "Squaring the Circle," vol. 5, pp. 132–133 (1913).
(32) "On the integral $\int_0^x \text{arc tan } t . \frac{dt}{t}$," vol. 7, pp. 93–96 (1915).
(33) "On the Divisors of a Number," vol. 7, pp. 131–134 (1915).
(34) "The Sum of the Square Roots of the First n Natural Numbers," vol. 7, pp. 173–175 (1915).
(35) "On the Product $\pi\left[1+\frac{x^2}{(a+nd)^2}\right]$," vol. 7, pp. 209–212 (1915).
(36) "Some Definite Integrals," vol. 11, pp. 81–88 (1919).
(37) "A Proof of Bertrand's Postulate," vol. 11, pp. 181–183 (1919).
(38) (Communicated by S. Narayana Aiyar), vol. 3, p. 60 (1911).

(B) Questions proposed and solved.

Nos. 260, 261, 283, 289, 294, 295, 298, 308, 353, 358, 386, 427, 441, 464, 489, 507, 541, 546, 571, 605, 606, 629, 642, 666, 682, 700, 723, 724, 739, 740, 753, 768, 769, 783, 785.

(C) Questions proposed but not solved as yet.

Nos. 284, 327, 359, 387, 441, 463, 469, 524, 525, 526, 584, 661, 662, 681, 699, 722, 738, 754, 770, 784, 1049, 1070, and 1076.

Finally, I may mention the following writings by other authors, concerned with Ramanujan's work :—

"Proof of a Formula of Mr. Ramanujan," by G. H. Hardy, 'Messenger of Mathematics,' vol. 44, pp. 18–21 (1915).
"Mr. S. Ramanujan's Mathematical Work in England," by G. H. Hardy (Report to the University of Madras, 1916. Privately printed).
"On Mr. Ramanujan's Empirical Expansions of Modular Functions," by L. J. Mordell, 'Proc. Camb. Phil. Soc.,' vol. 19, pp. 117–124 (1917).
"On Mr. Ramanujan's Congruence Properties of $p(n)$," by H. B. C. Darling, 'Proc. Cambridge Phil. Soc.,' vol. 19, pp. 217–218 (1919).

---

† All of Ramanujan's manuscripts passed through my hands, and I edited them very carefully for publication. The earlier ones I rewrote completely. I had no share of any kind in the results, except of course when I was actually a collaborator, or when explicit acknowledgment is made. Ramanujan was almost absurdly scrupulous in his desire to acknowledge the slightest help.

"Life Sketch of Ramanujan," editorial in the 'Journal of the Indian Math. Soc.,' vol. 11, p. 122 (1919).

"Note on the Parity of the Number which Enumerates the Partitions of a Number," by P. A. MacMahon, 'Proc. Cambridge Phil. Soc.,' vol. 20, pp. 281-283 (1921).

"Proof of certain Identities and Congruences Enunciated by S. Ramanujan," by H. B. C. Darling, 'Proc. London Math. Soc.,' ser. 2, vol. 19, pp. 350-372 (1921).

"On a Type of Modular Relation," by L. J. Rogers, *ibid.*, pp. 387-397.

It is plainly impossible for me, within the limits of a notice such as this, to attempt a reasoned estimate of Ramanujan's work. Some of it is very intimately connected with my own, and my verdict could not be impartial; there is much, too, that I am hardly competent to judge; and there is a mass of unpublished material, in part new and in part anticipated, in part proved and in part only conjectured, that still awaits analysis. But it may be useful if I state, shortly and dogmatically, what seems to me Ramanujan's finest, most independent, and most characteristic work.

His most remarkable papers appear to me to be (3), (7), (9), (17), (18), (19), and (21). The first of these is mainly Indian work, done before he came to England; and much of it had been anticipated. But there is much that is new, and, in particular, a very remarkable series of algebraic approximations to π. I may mention only the formulæ

$$\pi = \frac{63}{25}\frac{17+15\sqrt{5}}{7+15\sqrt{5}}, \qquad \frac{1}{2\pi\sqrt{2}} = \frac{1103}{99^2},$$

correct to 9 and 8 places of decimals respectively.

The long memoir (7) represents work, perhaps, in a backwater of mathematics, and is somewhat overloaded with detail; but the elementary analysis of "highly composite" numbers—numbers which have more divisors than any preceding number—is exceedingly remarkable, and shows very clearly Ramanujan's extraordinary mastery over the algebra of inequalities. Papers (9) and (17) should be read together, and in connection with Mr. Mordell's paper mentioned above; for Mr. Mordell afterwards proved a great deal that Ramanujan conjectured. They contain, in particular, very beautiful contributions to the theory of the representation of numbers by sums of squares. But I am inclined to think that it was in the theory of partitions, and the allied parts of the theories of elliptic functions and continued fractions, that Ramanujan shows at his very best. It is in papers (18), (19), and (21), and in the papers of Prof. Rogers and Mr. Darling, that I have quoted, that this side of his work (so far as it has been published) is to be found. It would be difficult to find more beautiful formulæ than the "Rogers-Ramanujan" identities, proved in (19); but here Ramanujan must take second place to Prof. Rogers; and, if I had to select one formula from all Ramanujan's work, I would agree with Major MacMahon in selecting a formula from (18), viz.,

$$p(4)+p(9)x+p(14)x^2+\ldots = 5\frac{\{(1-x^5)(1-x^{10})(1-x^{15})\ldots\}^5}{\{(1-x)(1-x^2)(1-x^3)\ldots\}^6},$$

where $p(n)$ is the number of partitions of n.

I have often been asked whether Ramanujan had any special secret; whether his methods differed in kind from those of other mathematicians; whether there was anything really abnormal in his mode of thought. I cannot answer these questions with any confidence or conviction; but I do not believe it. My belief is that all mathematicians think, at bottom, in the same kind of way, and that Ramanujan was no exception. He had, of course, an extraordinary memory. He could remember the idiosyncrasies of numbers in an almost uncanny way. It was Mr. Littlewood (I believe) who remarked that "every positive integer was one of his personal friends." I remember once going to see him when he was lying ill at Putney. I had ridden in taxicab No. 1729, and remarked that the number (7.13.19) seemed to me rather a dull one, and that I hoped it was not an unfavourable omen. "No," he replied, "it is a very interesting number; it is the smallest number expressible as a sum of two cubes in two different ways." I asked him, naturally, whether he knew the answer to the corresponding problem for fourth powers; and he replied, after a moment's thought, that he could see no obvious example, and thought that the first such number must be very large. His memory, and his powers of calculation, were very unusual, but they could not reasonably be called "abnormal." If he had to multiply two large numbers, he multiplied them in the ordinary way; he would do it with unusual rapidity and accuracy, but not more rapidly or more accurately than any mathematician who is naturally quick and has the habit of computation. There is a table of partitions at the end of our paper (15). This was, for the most part, calculated independently by Ramanujan and Major MacMahon; and Major MacMahon was, in general, slightly the quicker and more accurate of the two.

It was his insight into algebraical formulæ, transformations of infinite series, and so forth, that was most amazing. On this side most certainly I have never met his equal, and I can compare him only with Euler or Jacobi. He worked, far more than the majority of modern mathematicians, by induction from numerical examples; all of his congruence properties of partitions, for example, were discovered in this way. But with his memory, his patience, and his power of calculation, he combined also a power of generalisation, a feeling for form, and a capacity for rapid modification of his hypotheses, that was often really startling, and made him, in his own peculiar field, without a rival in his day.

It is often said that it is much more difficult now for a mathematician to be original than it was in the great days when the foundations of modern analysis were laid; and no doubt in a measure it is true. Opinions may differ as to the importance of Ramanujan's work, the kind of standard by which it should be judged, and the influence which it is likely to have on the mathematics of the future. It has not the simplicity and the inevitableness of the very greatest work; it would be greater if it were less strange. One gift it has which no one can deny, profound and invincible originality. He would probably have been a greater mathematician if he had been caught and tamed a little in his youth; he would have discovered more that was new, and that, no doubt, of greater importance. On the other hand, he would have been less of a Ramanujan, and more of a European professor, and the loss might have been greater than the gain.

G. H. H.

WOLDEMAR VOIGT, 1850—1919.

WOLDEMAR VOIGT was born at Leipzig in the year 1850. His school education was followed by attendance at the University there until his studies were interrupted by the outbreak of the Franco-German war of 1870, in which he served. His University course was afterwards completed at Königsberg in the years 1871–74. He here came under the influence of Franz Neumann, which determined largely the character of his own subsequent work. Alike in the type of subjects to which he devoted himself, in the formal elegance of his mathematical expositions, and in the severe precision of his style we can trace the inspiration of his illustrious teacher, for whom he retained a profound veneration. His abilities quickly gained recognition: he was made Extraordinary Professor at Königsberg in 1875, and was called to Göttingen as Ordinary Professor of Theoretical Physics in 1883. He held this post till his death on December 13, 1919.

The complexion of his scientific work is, on the whole, indicated by the title of his chair. It was, however, always in close relation to phenomena, and he carried out a classical series of measurements on the elasticity of crystals which demanded the utmost precision. For experimental work, except on a minute scale, he appears to have had, till a late period, scanty facilities. His favourite province, which he cultivated with a life-long enthusiasm, was the physical properties of crystals. In particular his determinations of elastic constants of various crystals may claim to have set at rest a historical controversy. The special form of molecular hypothesis on which the theories of Poisson and Cauchy were based involved the conclusion that certain relations must exist between the constants of a crystal whatever its classification. This would reduce the number of constants in a general scheme to fifteen. Voigt succeeded in proving that the relations in question were, in many cases, not even approximately fulfilled. This may be regarded as a vindication of the attitude of Green and his followers in this country who, avoiding special hypothesis, postulated only the principle of energy. Voigt himself did not accept this point of view as final, and indicated lines on which the old molecular hypothesis might be amended. A masterly review of the whole subject of crystalline elasticity was communicated by him to the Physical Congress held in Paris in 1900. This is distinguished by its orderly arrangement, and by the extreme elegance of the mathematical developments. Like Maxwell, he attached great importance to the classification and nomenclature of mathematical concepts; it is to him in particular that we owe the term "tensor," whose application has recently been so much extended. Voigt was also keenly interested in the other physical properties of crystals, and published an extensive treatise on 'Kristallphysik' (1910). He was the author of various papers on optical and electrical subjects, and of a systematic treatise on 'Magneto- und Elektro-optik' (1908). The breadth of his scientific sympathies may be further indicated by an allusion to his papers on Vortex Motion and on the Zeemann effect. In a paper on the Doppler effect published in 1887, Voigt was the first to establish the electromagnetic equations of transformation applicate to moving systems, which have played so important a part in the initial stages of the principle of relativity. A text-book in two volumes on Thermodynamics was dedicated to Lord Kelvin.

Voigt was a frequent visitor to this country, often as the delegate of the Göttingen Academy. He was present, for instance, at the Stokes jubilee in 1899, at the Owens College jubilee in 1902, at the Cambridge meeting of the British Association in 1904, at the St. Andrews 500th anniversary in 1911, and at the Royal Society celebration in 1912. No more acceptable representative could have been chosen. He had a generous appreciation of this country and its institutions, and an especial admiration of its leading physicists, in particular Stokes and Kelvin. He received many honorary degrees from British Universities, and was elected a Foreign Member of the Royal Society in 1913.

His many friends in this country felt an especial pang when the war came to interrupt intercourse with one whom they had found to be so sympathetic. It is to be recorded that even after the outbreak he made a courageous protest against the indiscriminate disparagement of everything English or French which was rife in Germany. He ventured to assert that there were elements in French and English culture which his own countrymen would do well at least to respect. On the general merits of the struggle he adopted the national view, which became embittered as the war went on. Those who knew him on his visits here will prefer to forget the alienation of the later years, and to recall only the grave and dignified courtesy, the friendly recognition, and the single-minded devotion to scientific truth which marked a noble and loveable personality.

H. L.

INDEX TO VOL. XCIX. (A)

Abernethy (C. L.) See Garner and Abernethy.

Abney (Sir W. de W.) Obituary notice of, i.

Adam (N. K.) The Properties and Molecular Structure of Thin Films of Palmitic Acid on Water. Part I, 336.

Alloys of aluminium, copper, and zinc (Haughton and Bingham), 47.

Alloys of iron with nickel, &c., influence of low temperatures on magnetic properties of (Onnes, etc.), 174.

Alpha-ray tracks, branched (Shimizu), 432.

Aluminium alloys (Haughton and Bingham), 47.

Armstrong (E. F.) and Hilditch (T. P.) A Study of Catalytic Actions at Solid Surfaces. VI.—Surface Area and Specific Nature of a Catalyst: two Independent Factors controlling the Resultant Activity, 490.

Aston (F. W.) On Isotopes (Discussion), 95.

Atomic lattices and electro-crystalline properties (Larmor), 1.

Balls (W. L.) A Simple Apparatus for Approximate Harmonic Analysis and for Periodicity Measurements, 283.

Beta-rays excited by γ-rays, magnetic spectrum of (Ellis), 261.

Bingham (K. E.) See Haughton and Bingham.

Bone (W. A.) Researches upon Brown Coals and Lignites. Part I.—Heat Treatment at Temperatures below 400° C. as a Possible Method for Enhancing their Fuel Values, 236.

Bowen (E. J.) See Hinshelwood and Bowen.

Brown coals and lignites, researches on, I (Bone), 236.

Brunt (D.) The Dynamics of Revolving Fluid on a Rotating Earth, 397.

Carbonic acid, total heat of liquid (Jenkin), 352.

Catalytic actions at solid surfaces. VI (Armstrong and Hilditch), 490.

Catalytic activity of copper (Palmer), 412.

Catalytic dehydrogenation of alcohols (Rideal), 153.

Catalytic reduction of ethylene to ethane (Palmer), 402.

Chakravarty (B. N.) On the Diffraction of Light incident at nearly the Critical Angle on the Boundary between two Media, 503.

Clifton (R. B.) Obituary notice of, vi.

Cyanogen, viscosity and molecular dimensions of (Rankine), 331.

Datta (S.) The Vacuum Arc Spectra of Sodium and Potassium, 69; —— On the Spectra of the Alkaline Earth Fluorides and their Relation to each other, 436.

Decomposition of crystalline solids, influence of physical conditions on (Hinshelwood and Bowen), 203.

Dehydrogenation, catalytic, of alcohols (Rideal), 153.

Diaphragms capable of continuous tuning (King), 163.

Discussion on isotopes, 87.

Dobbie (Sir J. J.) and Fox (J. J.) The Absorption of Light by Elements in a State of Vapour: the Halogens, 456.

Dynamics of revolving fluid on rotating earth (Brunt), 397.

Earth's crust, age of (Russell), 84.
Eddington (A. S.) A Generalisation of Weyl's Theory of the Electromagnetic and Gravitational Fields, 104.
Einstein's theory of gravitation (Jeffery), 123.
Electro-crystalline properties conditioned by atomic lattices (Larmor), 1.
Electromagnetic and gravitational fields, generalisation of Weyl's theory (Eddington), 104.
Electron, field of, on Einstein's theory of gravitation (Jeffery), 123.
Ellis (C. D.) The Magnetic Spectrum of the β-Rays excited by γ-Rays, 261.
Evaporation from large expanses of water (Giblett), 472.

Films of palmitic acid on water (Adam), 336.
Fletcher (Sir L.) Obituary notice of, ix.
Fluid, dynamics of revolving, on rotating earth (Brunt), 397.
Fluorescence, interval between absorption and emission of light in (Wood), 362.
Fourier's integral theorem, extension of, and physical applications (Fowler), 462.
Fowler (R. H.) A Simple Extension of Fourier's Integral Theorem and some Physical Applications, in particular to the Theory of Quanta, 462.
Fox (J. J.) See Dobbie and Fox.
Functions of positive type and linear transformations (Mercer), 19.

Galvanometer records, method of analysing (Hartree and Hill), 172.
Garner (W. E.) and Abernethy (C. L.) Heats of Combustion and Formation of Nitro-Compounds. Part I.—Benzene, Toluene, Phenol and Methylaniline Series, 213.
Giblett (M. A.) Some Problems connected with Evaporation from Large Expanses of Water, 472.
Griffiths (E. A.) A Liquid Oxygen Vaporiser, 281.

Hadfield (Sir R. A.) See Onnes (and others).
Harmonic analysis, apparatus for approximate, and for periodicity measurements (Balls), 283.
Hartree (W.) and Hill (A. V.) A Method of Analysing Galvanometer Records, 172.
Hatschek (E.) Anomalous Liesegang Stratifications produced by the Action of Light, 496.
Haughton (J. L.) and Bingham (K. E.) The Constitution of the Alloys of Aluminium, Copper, and Zinc containing High Percentages of Zinc, 47.
Hilditch (T. P.) See Armstrong and Hilditch.
Hill (A. V.) See Hartree and Hill.
Hinshelwood (C. N.) and Bowen (E. J.) The Influence of Physical Conditions on the Velocity of Decomposition of certain Crystalline Solids, 203.

Ibbs (T. L.) Some Experiments on Thermal Diffusion, 385.
Integrals, transformation of (Young), 252.
Invariants, differential, and multenions (McAulay), 292.
Ionising rays, apparatus for detecting (Shimizu), 425.
Iron, influence of low temperatures on magnetic properties of alloys of (Onnes, etc.), 174.
Isotopes, discussion on, 87.

Jeffery (G. B.) The Field of an Electron on Einstein's Theory of Gravitation, 123.
Jenkin (C. F.) The Total Heat of Liquid Carbonic Acid, 352.

Kikuchi (T.) On the Moving Striations in a Neon Tube, 257.
King (L. V.) On the Design of Diaphragms capable of Continuous Tuning, 163.

Lamb (H.) and Southwell (R. V.) The Vibrations of a Spinning Disk, 272.
Larmor (Sir J.) On Electro-crystalline Properties as conditioned by Atomic Lattices, 1.
Liesegang stratifications, anomalous (Hatschek), 496.
Light, absorption by elements in state of vapour (Dobbie and Fox), 456.
Light, action of, on Liesegang stratifications (Hatschek), 496.
Light from night sky, colour of (Rayleigh), 10.
Light in fluorescence, absorption and emission of (Wood), 362.
Lignites and brown coals, researches on, I (Bone), 236.
Lindemann (F. A.) on Isotopes (Discussion), 102.
Linear transformations and functions of positive type (Mercer), 19.

McAulay (A.) Multenions and Differential Invariants, 292.
Magnetic properties of alloys of iron, influence of low temperatures on (Onnes, etc.), 174.
Mercer (J.) Linear Transformations and Functions of Positive Type, 19.
Merton (T. R.) On Spectrophotometry in the Visible and Ultra-violet Spectrum, 78; —— On Isotopes (Discussion), 101.
Multenions and differential invariants (McAulay), 292.

Neon tube, moving striations in (Kikuchi), 257.
Newman (F. H.) A New Form of Wehnelt Interrupter, 324.
Nitro-compounds, heats of combustion and formation of (Garner and Abernethy), 213.

Obituary Notices :—
  Abney, Sir W. de W., i.
  Clifton, R. B., vi.
  Fletcher, Sir L., ix.
  Ramanujan, S., xiii.
  Voigt, W., xxix.
Ocean currents, dissipation of energy in, relations with salinities and temperatures (Street), 39.
Onnes (K.), Hadfield (Sir R. A.), and Woltjer (H. R.) On the Influence of Low Temperatures on the Magnetic Properties of Alloys of Iron with Nickel and Manganese, 174.
Oxygen, liquid, vaporiser (Griffiths), 281.

Palmer (D. M.) and (W. G.) Some Experiments on the Catalytic Reduction of Ethylene to Ethane, 402.
Palmer (W. G.) The Catalytic Activity of Copper, Part II, 412.
Periodicity measurements, simple apparatus for (Balls), 283.
Phosphorus, study of the glow of (Rayleigh), 372.

Quanta, theory of, and Fourier's integral theorem (Fowler), 462.

Ramanujan (Srinivasa). Obituary notice of, xiii.
Rankine (A. O.) On the Viscosity and Molecular Dimensions of Gaseous Cyanogen, 331.
Rayleigh (Lord). The Colour of the Light from the Night Sky, 10; —— A Study of the Glow of Phosphorus.—Periodic Luminosity and Action of Inhibiting Substances, 372.
Rideal (E. K.) On the Catalytic Dehydrogenation of Alcohols, 153.
Rosenhain, W. The Hardness of Solid Solutions, 196.
Russell (H. N.) A Superior Limit to the Age of the Earth's Crust, 84.

Saha (M. N.) On a Physical Theory of Stellar Spectra, 135.
Shimizu (T.) A Reciprocating Expansion Apparatus for Detecting Ionising Rays, 425; —— A Preliminary Note on Branched α-Ray Tracks, 432.
Sky, colour of light from night (Rayleigh), 10.

Soddy (F.) on Isotopes (Discussion), 97.
Solid solutions, hardness of (Rosenhain), 196.
Southwell (R. V.) See Lamb and Southwell.
Spectra, arc, of sodium and potassium (Datta), 69; —— of alkaline earth fluorides (Datta), 436; physical theory of stellar (Saha), 135.
Spectrophotometry in visible and ultra-violet spectrum (Merton), 78.
Spectrum, magnetic, of β-rays excited by γ-rays (Ellis), 261.
Spinning disk, vibrations of (Lamb and Southwell), 272.
Street (R. O.) The Dissipation of Energy in Permanent Ocean Currents, with some Relations between Salinities, Temperatures and Currents, 39.

Thermal diffusion, experiments on (Ibbs), 385.
Thomson (Sir J. J.) on Isotopes (Discussion), 87.
Transformation of integrals (Young), 252.

Voigt (W.) Obituary notice of, xxix.

Wehnelt interrupter, new form of (Newman), 324.
Weyl's theory of electromagnetic field generalised (Eddington), 104.
Woltjer (H. R.) See Onnes (and others).
Wood (R. W.) The Time Interval between Absorption and Emission of Light in Fluorescence, 362.

Young (W. H.) On the Transformation of Integrals, 252.

END OF THE NINETY-NINTH VOLUME (SERIES A).

ERRATUM.

Vol. A 98, p. xlvii (Rayleigh Obituary Notice)
Line 18 from top, *for* Buffon *read* Boileau.

Harrison and Sons, Ltd., Printers in Ordinary to His Majesty, St. Martin's Lane.

I. A. R. I. 75

IMPERIAL AGRICULTURAL RESEARCH
INSTITUTE LIBRARY
NEW DELHI.

Date of issue.	Date of issue.	Date of issue.
23.[illegible].50		